Reducing Pyrrolysine tRNA Copy Number Improves the Performance of Genetic Code Expansion in Live Cell Imaging of Bioorthogonally Labeled Proteins

Noa Aloush,[†] Tomer Schvartz,[‡] Andres I. König,[‡] Sarit Cohen,[†] Dikla Nachmias,[‡] Oshrit Ben-David,[†] Natalie Elia,[‡] and Eyal Arbely*[,†,‡]

†*Department of Chemistry and National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva, 8410501, Israel*

‡*Department of Life Sciences, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva, 8410501, Israel*

E-mail: arbely@bgu.ac.il

Phone: +972-8-6428739. Fax: +972-8-6428449

Abstract

Genetic code expansion technology enables the incorporation of non-canonical amino acids (NCAAs) into proteins expressed in live cells. The NCAA is commonly encoded by an in-frame amber stop codon (TAG) and the methodology relies on the use of an orthogonal aminoacyl tRNA synthetase and its cognate amber suppressor tRNA; e.g., the pyrrolysine synthetase / $tRNA^{Pyl}_{CUA}$ (PylT) pair. It is widely accepted that in cultured mammalian cells, intracellular concentration of amber suppressor pyrrolysine tRNA is a limiting factor in amber-suppression efficiency. Therefore, multiple copies of *pylT* are usually encoded in current expression systems in order to improve NCAA

incorporation level. We quantified protein expression levels as a function of encoded *pylT* copy number, incorporated amino acid, and cell line. We found that a decrease in *pylT* copy number does not always correlate with a decrease in protein expression. Importantly, we found that reducing *pylT* copy number improved live-cell high-resolution imaging of bioorthogonally-labeled intracellular proteins by enhancing signal-to-noise ratio without affecting protein expression levels. This enabled us to label the intracellular layer of the plasma membrane as well as to co-label two proteins in a cell. Our results indicate that the number of encoded *pylT* genes should be carefully optimized based on the cell line, incorporated non-canonical amino acid, and the application it is used for.

Abbreviations

BCN-Lys, bicyclo[6.1.0]nonyne-L-lysine; BCN-RS, BCN-specific evolved synthetase; Boc-Lys, Nε-[(*tert*-butoxy)carbonyl]-L-lysine; EF1α, human elongation factor 1 α-subunit promoter; FLIP, fluorescence loss in photobleaching; GMFI geometric mean fluorescence intensity; NCAA, non-canonical amino acid; Pol III, RNA polymerase III; Pyl-RS, pyrrolysyl tRNA synthetae; SIM, structured illumination microscopy; SiR, silicon rhodamine

Keywords

bioorthogonal labeling, amber suppressor tRNA, super resolution microscopy

Introduction

Methods of genetic code expansion enable the site-specific incorporation of dozens of non-canonical amino acids (NCAAs), each with unique functionality, into proteins expressed in bacteria,[1] eukaryotes,[2–5] plants,[6] and multicellular organisms.[7–10] Current popular method-

ologies are based on an orthogonal aminoacyl-tRNA synthetase/tRNA pair that can facilitate the co-translational incorporation of a NCAA into the protein of interest in response to a specific codon; typically the amber stop codon, UAG.[11–13] Early successful attempts of NCAA incorporation into proteins expressed in cultured mammalian cells utilized orthogonal aminoacyl-tRNA synthetase/tRNA pairs of bacterial origin, usually *Escherichia coli* (*E. coli*) tyrosyl tRNA synthetase and either *Bacillus stearothermophilus* or *E. coli* $\mathrm{tRNA}^{\mathrm{Tyr}}$.[2,3,14,15] In recent years, the pyrrolysyl tRNA synthetase (Pyl-RS) and its cognate amber suppressor $\mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$ (PylT)[16,17] became one of the orthogonal pairs commonly used in cultured mammalian cells.[4,18] The Pyl-RS/PylT couple was found to be orthogonal in bacteria and eukaryotic cells, facilitating the evolution of orthogonal NCAA-specific aminoacyl synthetases. Additionally, Pyl-RS shows high substrate promiscuity and low selectivity towards the tRNA anticodon sequence, making this pair an attractive system for genetic code expansion.[18]

Initial, as well as recent, optimization and calibration protocols for expanding the genetic code of cultured mammalian cells were commonly performed using only one cell line, the human embryonic kidney (HEK293) cell line. More than that, the studied experimental systems were based on different orthogonal aminoacyl-tRNA synthetase/tRNA pairs, promoters, and terminators. In addition, the number of encoded tRNA genes, the number of plasmids, and the DNA delivery methods varied, making it difficult to compare between different experiments.[2–4,19–28] Despite the limited variety of cell lines and the large variability in experimental conditions, these studies provided some of the basic principles of current methodologies. It was concluded that the intracellular concentration of amber suppressor tRNA is a limiting factor in amber suppression efficiency and hence, in overall protein expression level. Moreover, it was found that high levels of tRNA transcription and processing can be achieved by using constitutive RNA polymerase III (Pol III) promoters such as U6 or H1 promoters, that have no downstream transcriptional elements.[3,4,20,22,24] Consequently, in the majority of current systems for genetic code expansion in cultured mammalian cells,

multiple copies of a U6 or H1 promoter followed by an amber suppressor tRNA, are encoded in tandem and/or on different plasmids.

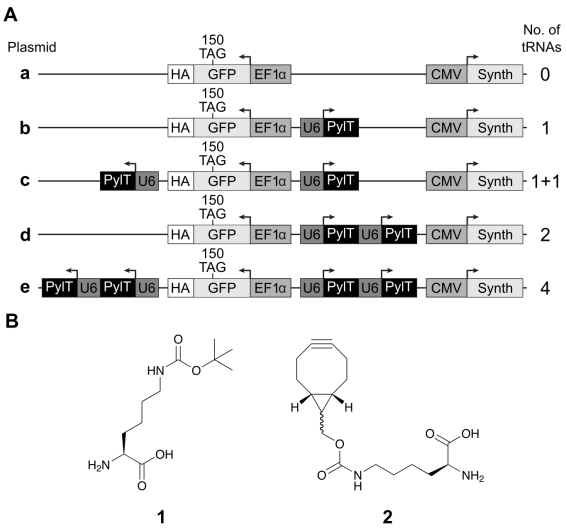

Figure 1: A) Schematic representation of plasmids carrying different number of encoded *pylT* genes and $EGFP^{150TAG}$-*HA*. Plasmid **a**, with no encoded tRNA was used as control and plasmid **b** carries 1 copy of *pylT*. Plasmids **c** and **d** carry two encoded copies of *pylT* cloned separately (plasmid **c**, marked '1+1') or in tandem (plasmid **d**, marked '2'). Plasmid **e** carries 4 copies of *pylT* in two cassettes of two *pylT* genes, cloned in tandem. One set of plasmids was cloned with wild-type Pyl-RS and a second set was cloned with evolved BCN-RS (total of 10 plasmids). B) Chemical structures of amino acids used in the current study: Boc-Lys (**1**) and BCN-Lys (**2**).

Proper balance between a given tRNA and its cognate aminoacyl-tRNA synthetase is important for maintaining accurate and efficient aminoacylation, as well as for high amber suppression efficiency.[22,29] Nevertheless, there are very few studies in which amber suppression efficiency in cultured mammalian cells was measured as a function of the ratio between the number of encoded *pylT* genes and the cognate pyrrolysyl-tRNA synthetase.[26,27] A systematic and quantitative study has yet to be performed with the aim of identifying the minimal number of encoded *pylT* genes required for efficient amber suppression in different cell lines as function of the encoded NCAA.

Results and discussion

We have recently demonstrated the efficiency of a single-plasmid–based system for the genetic incorporation of NCAAs into proteins expressed in transiently transfected cultured mammalian cells.[30] In contrast to transient co-transfection with two or more plasmids, where different populations of cells may take up different combinations of plasmid DNA,[31,32] encoding all the genetic components on one plasmid allows for a better control over the ratio between different genes introduced into the transfected cells. Taking advantage of this expression system, we created a set of plasmids for expression of $EGFP^{150TAG}$-*HA* controlled by the human elongation factor 1 α-subunit promoter (EF1α, plasmids **a**–**e**, Figure 1A). These plasmids include 0 to 4 copies of the U25C mutant of *Methanosarcina mazei* (*M. mazei*) *pylT* gene controlled by constitutive U6 promoter.[33] To allow direct comparison between the expression of proteins carrying different NCAAs, plasmids **a**–**e** were cloned with wild-type Pyl-RS for the incorporation of Nε-[(*tert*-butoxy)carbonyl]-L-lysine (Boc-Lys, **1**, Figure 1B) as well as with an evolved synthetase (BCN-RS) for the incorporation of bicyclo[6.1.0]nonyne-L-lysine (BCN-Lys, **2**, Figure 1B). Expression of the synthetase gene is controlled by the human cytomegalovirus (CMV) immediate-early promoter.

With the above described plasmids in hand, we estimated the incorporation level of NCAAs **1** and **2** into 150TAG-mutant of EGFP-HA expressed in HEK293T and COS7 cells by quantifying protein expression level using Western blot analyses (Figure 2). As previously reported by others,[19,24,27,28] amber suppression efficiency in HEK293T cells improved when the number of encoded *pylT* genes was increased from 1 to 4. However, the effect was more notable for BCN-RS, compared to Pyl-RS (Figure 2A). Interestingly, our data suggest that for HEK293T cells, cloning *pylT* genes in tandem is less efficient than cloning the same number of genes in separate locations on the plasmid. Expression of 150TAG-mutant of EGFP-HA was significantly higher when HEK293T cells were transfected with plasmid **c** (2 copies of *pylT* cloned separately) compared to plasmid **d** (2 copies of *pylT* cloned in tandem, Figure 2A).

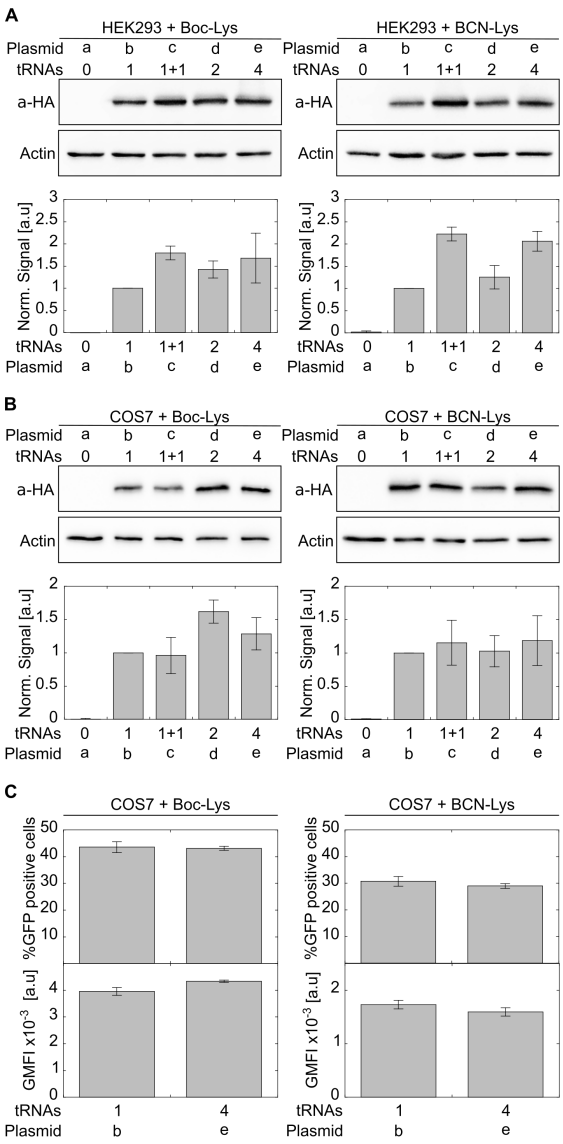

Figure 2: HEK293T (A) or COS7 cells (B) were transfected with plasmids **a**–**e** carrying either Pyl-RS for the incorporation of **1** or evolved BCN-RS for the incorporation of **2**. Expression of $EGFP^{150TAG}$-*HA* was quantified by Western blot using the C-terminus HA-tag, and normalized to actin. Expression levels as function of indicated tRNA copy-number are displayed relative to plasmid carrying 1×tRNA ± SSD (n=3 for HEK293T cells and n=5 for COS7 cells). C) COS7 cells were transiently transfected with plasmids carrying either Pyl-RS or BCN-RS and 1 or 4 copies of tRNA. Transfected cells incubated with indicated NCAA for 48 h and level of EGFP expression was analyzed by FACS. Percent of GFP-positive cells is presented as mean value ± SSD (n=3). Protein expression level was calculated from GFP fluorescence intensity (GFP-area) and displayed as geometric mean fluorescence intensity (GMFI ± SSD, n=3).

In contrast to HEK293T cells, in COS7 cells, reducing the number of encoded *pylT* genes from 4 to 1 had minor effect on the expression level of 150TAG-mutant of EGFP-HA (Figure 2B). This result was independent of the encoded NCAA, as it was observed when cells were transiently transfected with plasmids carrying *Pyl-RS* or *BCN-RS* in the presence of NCAA **1** or **2**, respectively. We therefore decided to focus on plasmids that carry the minimal and maximal number of *pylT* genes (1×PylT plasmid **b** and 4×PylT plasmid **e**, respectively) and analyzed the expression level of 150TAG-mutant of EGFP-HA expressed in transfected COS7 cells by flow cytometry (Figure 2C). Flow cytometry measurements showed that reducing the number of encoded *pylT* genes from 4 to 1 has no effect on the number of cells expressing full length EGFP-HA or protein expression level, in excellent agreement with Western blot analyses. The data in Figure 2 was collected 48 h post transfection with similar results measured 24 h post transfection (Supplementary Figure S1). Therefore, encoding multiple copies of *pylT* genes, in an attempt to increase intracellular levels of PylT, does not necessarily correlate with an increase in amber suppression efficiency and protein expression level. In HEK293T cells, amber suppression efficiency was affected by the number of *pylT* genes and their organization (cloned separately or in tandem), while the effect was NCAA (i.e., aminoacyl tRNA synthetase) dependent. However, in COS7 cells, amber suppression efficiency was nearly independent of the number of encoded *pylT* genes.

One of the exciting applications of genetic code expansion technology is the site specific incorporation of NCAAs carrying a bioorthogonal functional group for chemoselective labeling with fluorescent organic dyes, within living cells.[34–43] The site-specific labeling of proteins with organic fluorophores offers a superior alternative to the commonly used fluorescent proteins in fluorescent imaging of live cells.[35,44] Not only is such labeling not limited to the N- or C-terminus, but also organic fluorophores are much smaller than fluorescent proteins (∼0.5 kDa *vs.* ∼27 kDa) and therefore may have less effect on structure and cellular function of the labeled protein. In addition, organic fluorophores often have excellent photophysical properties which may allow for prolonged live cell fluorescent imaging with reduced photo-

toxicity. In recent years, the site specific labeling of proteins *via* inverse-electron-demand Diels-Alder reaction has been demonstrated with different NCAAs carrying strained systems (i.e. alkene or alkyne) and tetrazine-conjugated organic fluorophores.[45] For example, proteins carrying NCAA **2** can be labeled in live cells with 1,2,4,5-tetrazine-conjugated fluorophores and visualized by fluorescent microscopy techniques.[37–39,41]

A current challenge in live cell imaging of bioorthogonally labeled intracellular proteins is high background fluorescence observed upon the addition of cell-permeable tetrazine-conjugated fluorophores.[41,43] One of the sources for this background fluorescence (apart from incorporation of the NCAA in response to endogenous amber codons) is the reaction between tetrazine-conjugated fluorophores and aminoacylated tRNAs that forms labeled tRNAs.[41,43] While small, cell permeable fluorophores can be readily washed out of cells, labeled tRNAs are not cell-permeable and may remain in cells for hours. Our data suggest that in transiently transfected COS7 cells, only one copy of encoded *pylT* gene is sufficient for efficient incorporation of NCAA **2** into expressed proteins. Therefore, we decided to measure the effect of the number of encoded *pylT* genes on NCAA incorporation in the context of bioorthogonal labeling and live cell imaging. We designed four plasmids to allow for expressing and visualizing fluorescently labeled proteins in live cultured mammalian cells as a function of encoded *pylT* genes (plasmids **f**–**i**, Figure 3A). As a model for membrane protein labeling, we cloned $EGFP^{150TAG}$-*CAAX* for the expression of EGFP-CAAX anchored to the cytosolic side of plasma membrane *via* prenylation of the C-terminus CAAX motif. In addition, we cloned *HA-α-tubulin*^{46TAG} as a model for cytosolic protein labeling, with the aim of labeling cellular microtubules. As a fluorophore for bioorthogonal labeling we chose the cell-permeable tetrazine-conjugated silicon rhodamine (SiR-Tet) fluorophore (Supplementary Figure S2).

COS7 cells were transfected in the presence of NCAA **2** with 1×PylT plasmid **f** or **h**, or with 4×PylT plasmid **g** or **i** (Figure 3A). Cells were labeled with SiR-Tet and visualized by live confocal microscopy. In cells transfected with plasmid **f** or **g**, green fluorescence labeling

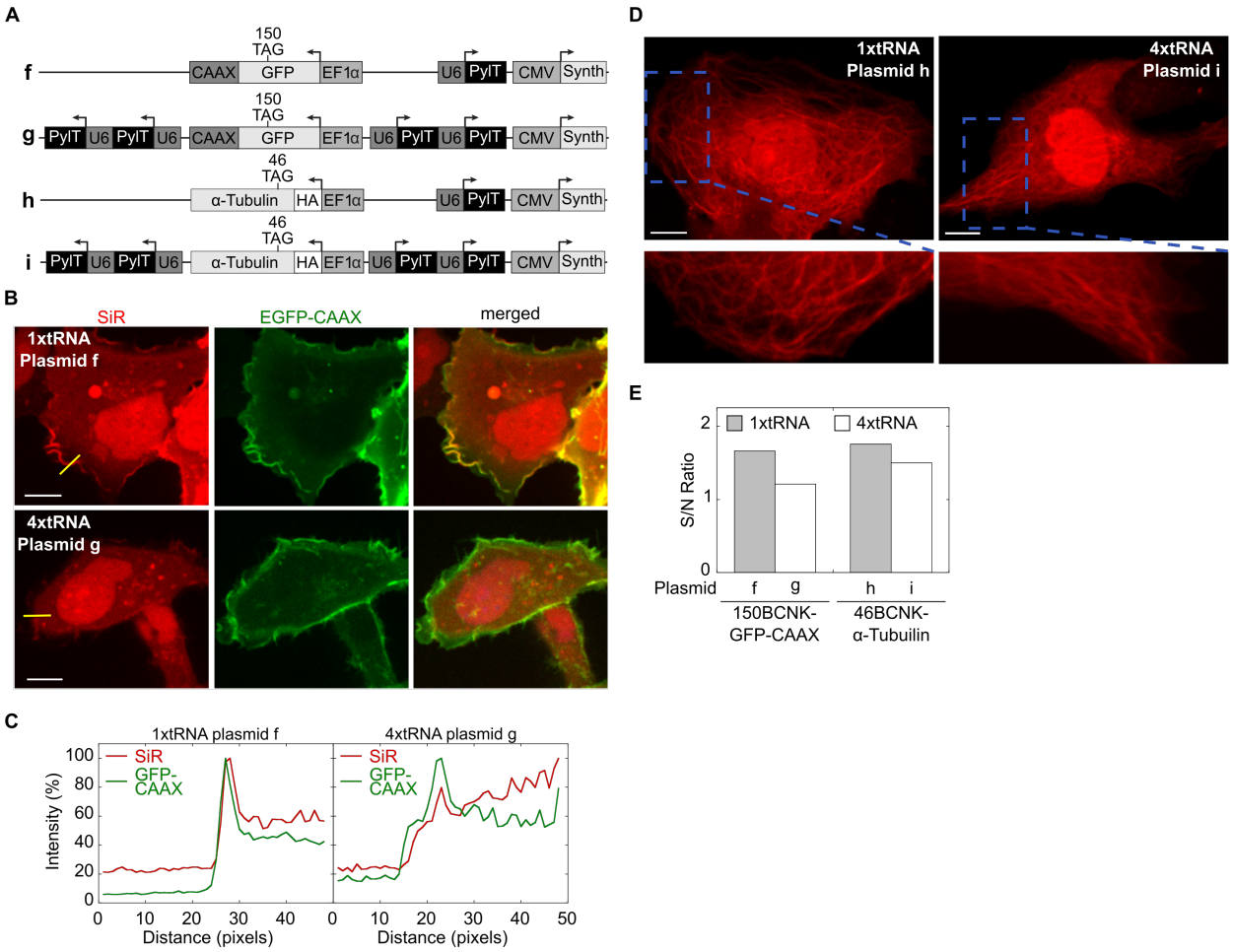

Figure 3: A) Schematic representation of 1×PylT or 4×PylT plasmids, for the expression of EGFP-CAAX (plasmids **f**, **g**) or α-tubulin (plasmids **h**, **i**) with site-specifically incorporated NCAA **2**. B) Live cell imaging of SiR-labeled 150BCNK-EGFP-CAAX expressed in COS7 cells. Left panel: red SiR; center panel: green EGFP; right panel: merge. Scale bar 10 µm. C) Line intensity profiles of 488 and 640 nm channels (EGFP and SiR, respectively) are presented as percent of maximum value and plotted as function of distance along a line (yellow line in left panel B). D) Live cell imaging of SiR-labeled 46BCNK-α-tubulin expressed in COS7 cells. Zoomed-in images of a subset of the cell are presented bellow (rotated by 90°). Scale bar 10 µm. E) Signal-to-noise ratios calculated from live-cell imaging of SiR-labeled 150BCNK-EGFP-CAAX and 46BCNK-α-tubulin expressed in COS7 cells using plasmids carrying 1 or 4 copies of *pylT*.

of the membrane was visible when either plasmid was used (Figure 3B). Consistent with our previous observations (Figure 2B and C), these results confirm that expression level and membrane localization of 150BCNK-EGFP-CAAX were independent of *pylT* copy number. The number of encoded *pylT* genes, however, had a significant effect on imaging of 150BCNK-

EGFP-CAAX using the SiR-Tet label (red panel). When cells were transfected with 4×PylT plasmid **g**, specific SiR-labeling of membrane-anchored 150BCNK-EGFP-CAAX could not be detected. In contrast, when 150BCNK-EGFP-CAAX was expressed using 1×PylT plasmid **f**, SiR-labeling of membrane-anchored 150BCNK-EGFP-CAAX was visible. To evaluate whether these differences could stem from differences in background fluorescence, we measured EGFP-CAAX and SiR fluorescence intensities along a line that crosses the plasma membrane (yellow line, Figure 3B, left panel). While a sharp peak in intensity levels was obtained in the GFP channel, representing membrane-anchored 150BCNK-EGFP-CAAX, practically no peak in intensity was obtained in SiR channel for 150BCNK-EGFP-CAAX expressed from 4×PylT plasmid **g**. Instead, SiR fluorescence intensity levels measured in the cytosol were similar or even lower than fluorescence intensity levels measured in the membrane (Figure 3C). However, when 150BCNK-EGFP-CAAX was expressed from 1×PylT plasmid **f**, a sharp peak in intensity levels was obtained in both GFP and SiR channels, representing membrane-anchored 150BCNK-EGFP-CAAX (Figure 3C). Moreover, under these conditions, SiR fluorescence intensity levels were significantly higher in the membrane than in the cytosol, indicating that the ability to observe plasma membrane labeling using SiR fluorescence under these conditions results from improvement in the signal to noise ratio. It appears that signal to noise ratio was significantly improved upon reduction of the number of encoded pylT genes to one.

The use of 1 copy of encoded *pylT* gene, instead of 4, also improved image quality of SiR-labeled 46BCNK-α-tubulin expressed in COS7 cells (Figure 3D). In cells transfected with 4×PylT plasmid **i**, few and relatively thick SiR-labeled fibers could be seen in only a small subset of the transfected cells. In contrast, SiR-labeled microtubules were clearly observed in cells transfected with 1×PylT plasmid **h**. Overall, the improvement in image quality of SiR-labeled 150BCNK-EGFP-CAAX and 46BCNK-α-tubulin, correlates with an increase of 37% and 17% in signal-to-noise ratio, respectively (Figure 3E).

The SiR-labeling found in the cytosol—and more significantly in the nucleus (Figure 3B

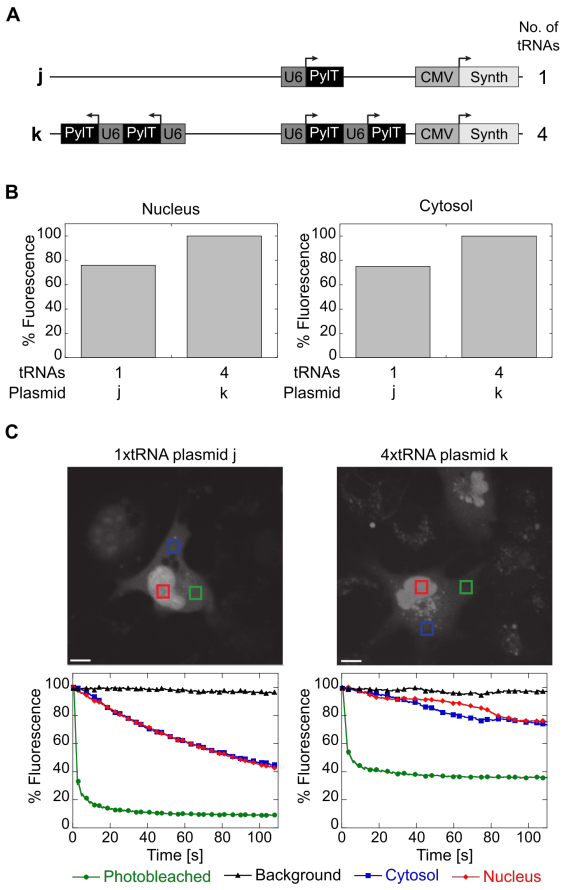

Figure 4: A) Schematic representation of plasmids with 1 or 4 copies of *pylT* genes and BCN-RS used for characterizing non-protein-specific labeling and dynamics of aminoacylated PylT. B) Relative average SiR-fluorescence measured in the nucleus (left) or cytosol (right) of COS7 cells expressing plasmid **j** or **k** and labeled with SiR-Tet. C) FLIP measurement of COS7 cells described in panel B. Mean fluorescence intensity was measured in the photobleached area (green), nearby cell (background, black) as well as in the cytosol and nucleus of the photobleached cell (blue and red, respectively) and displayed as function of time. Scale bar 10 µm.

and D)—was NCAA-dependent and therefore the result of a reaction between NCAA **2** and SiR-Tet (Supplementary Figure S3). As mentioned above, this chemoselective and non-protein-specific background labeling is believed to be caused by a reaction between tetrazine-conjugated fluorophores and aminoacylated tRNAs.[41,43] Consistent with this notion, we measured an increase in signal-to-noise ratio when the number of encoded *pylT* genes was reduced from 4 to 1 (Figure 3E). We therefore decided to characterize the effects of a reduced number of encoded *pylT* genes on *overall* non-protein-specific labeling. To do

so, we created plasmids with 1 or 4 copies of *pylT* genes and BCN-RS (plasmids **j** and **k**, Figure 4A). Lacking a protein-of-interest gene with an in frame TAG mutation, plasmids **j** and **k** allow for direct visualization and quantification of background fluorescence resulting from SiR–Tet-labeled PylT, without masking from the co-expressed and labeled protein. COS7 cells were transiently transfected with plasmid **j** or **k** in the presence of NCAA **2** and labeled with SiR-Tet. Live-cell confocal microscopy revealed that the averaged non-protein-specific fluorescence intensity in the nucleus and cytosol was reduced by 24% and 26%, respectively, when *pylT* copy number was reduced from 4 to 1 (Figure 4B). Hence, reducing the number of encoded *pylT* genes improved the signal-to-noise ratio by reducing background SiR-fluorescence. Furthermore, these results suggest that aminoacylated PylT is not entrapped at the nucleus, but rather at a state of dynamic equilibrium between the nucleus and cytosol. This equilibrium is independent of the number of encoded *pylT* genes, since the use of one *pylT* gene compared to four, reduced the population of aminoacylated PylT equally in the cytosol and the nucleus.

It was recently suggested that the fraction of nuclear PylT can be minimized by active nuclear export through the addition of a nuclear export signal to the evolved synthetase.[43] However, an increase in cytosolic fraction of PylT may interfere with live-cell imaging due to labeling of cytosolic and aminoacylated PylT. In addition, our understanding of PylT equilibrium between the nucleus and cytosol, as well as the factors that may affect it, is still limited. In order to study the dynamics of aminoacylated PylT between the nucleus and the cytosol we measured fluorescence loss in photobleaching (FLIP) in COS7 cells transfected with 1×PylT plasmid **j** or 4×PylT plasmid **k** and labeled with SiR-Tet (Figure 4C). Over the course of two minutes, for a given number of encoded *pylT* genes, the photobleaching pattern in the nucleus was similar to the one in the cytosol, suggesting that the nuclear fraction of aminoacylated PylT is in equilibrium with the cytosolic fraction. This result is consistent with the equal reduction in nuclear and cytosolic background fluorescence observed upon reduction in *pylT* copy number (Figure 4B). Although the equilibrium between the

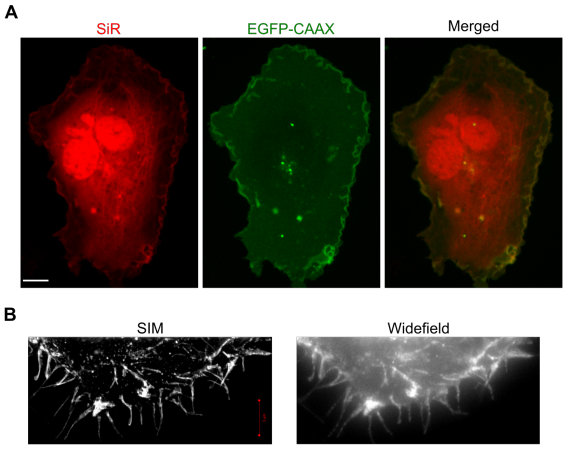

Figure 5: A) COS7 cells co-transfected with plasmids **f** and **h** in the presence of BCN-Lys and labeled with SiR-Tet. Using one copy of encoded *pylT* gene, enabled the co-expression and SiR labeling of membrane anchored 150BCNK-EGFP-CAAX and 46BCNK-α-tubulin. Scale bar 10 µm. B) Zoomed-in regions (rotated by 90 degrees) taken from SIM and widefield imaging of SiR-labeled 150BCNK-EGFP-CAAX (Supplementary Figure S4).

nucleus and cytosol was not affected by the number of encoded *pylT* genes, we measured a significant and consistent difference in overall photobleaching efficiency which was affected. In cells transfected with 1×PylT plasmid **j** the fluorescence in the photobleached area was reduced to less than 10% and fluorescence in the cytosol and nucleus was reduced to 40% after two minutes (Figure 4C, left panel). In contrast, in cells transfected with 4×PylT plasmid **k**, fluorescence in the bleached area was only reduced to 40% and fluorescence in the cytosol and nucleus was reduced to about 70% when the same FLIP protocol was used for two minutes (Figure 4C, right panel). These results suggest that in COS7 cells transfected with 4×PylT plasmid there is an immobile, photobleaching-resistant population of aminoacylated and SiR-labeled PylT. While this photobleaching-resistant population may increase the non-protein-specific background labeling, its origin and character have yet to be determined.

In light of the observed improvement in signal-to-noise ratio following the reduction in number of encoded *pylT* genes, we decided to use our single-plasmid–based system (that facilitates co-transfection) to co-label two different proteins for live-cell fluorescence imaging. COS7 cells were co-transfected with 1×PylT plasmids **f** *and* **h**, or 4×PylT plasmids **g** *and*

i (Figure 3A), for co-expression and labeling of 150BCNK-EGFP-CAAX *and* 46BCNK-α-tubulin. Co-labeling experiments using 4×PylT plasmids **g** and **i** were unsuccessful, as image quality was bellow the level obtained with single protein labeling. In contrast, live confocal microscopy showed clear SiR-labeling of both membrane anchored 150BCNK-EGFP-CAAX and 46BCNK-α-tubulin in COS7 cells co-transfected with 1×PylT plasmids **f** and **h** (Figure 5A). In addition, there was clear overlap between GFP fluorescence and SiR-labeling of the membrane. To the best of our knowledge, this is the first example of live-cell imaging of two different proteins bioorthogonally labeled with organic fluorophores *via* genetic encoding of NCAAs.

As a final testament to our conclusion that for imaging applications only one copy of *pylT* gene is required for incorporation of NCAA **2** into proteins expressed in COS7 cells, we visualized SiR-labeled 150BCNK-EGFP-CAAX using high-resolution structured illumination microscopy (SIM, Figure 5B). COS7 cells were transfected with 1×PylT plasmid **f** in the presence of NCAA **2**, labeled with SiR-Tet and fixed before visualized by SIM. Using these conditions, we were able to generate a 3D high-resolution image of the labeled membrane in SIM (Figure 5B and Supplementary Figure S4). This result demonstrates that even with challenging fluorescent imaging applications that require high labeling density, such as SIM, one copy of encoded *pylT* gene can allow for efficient protein expression level and subsequent bioorthogonal labeling of intracellular proteins.

In summary, we have demonstrated that the genetic encoding of multiple *pylT* genes does not necessarily improve the amber suppression efficiency when using the orthogonal pyrrolysine tRNA synthetase/$\mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$ pair in transiently transfected cultured mammalian cells. It has been suggested that amber suppression efficiency is limited by the intracellular concentration of PylT. We found that this notion is true in some cases (e.g., HEK293T cells), while in other cases, increasing *pylT* copy number has negligible effect on amber suppression efficiency and overall protein expression level (e.g., COS7 cells). Specifically, we presented significant improvement in labeling and image quality, as well as signal-to-noise ratio, in

live cell imaging of SiR-labeled proteins expressed in transiently transfected COS7 cells, by reducing the *pylT* copy number from 4 to 1. We suggest that for applications of genetic code expansion technology in cultured mammalian cells, the number of encoded *pylT* genes should be optimized based on the cell line and incorporated NCAA.

Acknowledgement

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 678461 (to E.A.) and No. 639313 (to N.E.), from the ARO under grant agreement No. 65422-LS (to N.E. and E.A.), and from the Israel Science Foundation (grant number 807/15 to E.A.).

References

1. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the Genetic Code of Escherichia coli. *Science 292*, 498–500.

2. Sakamoto, K., Hayashi, A., Sakamoto, A., Kiga, D., Nakayama, H., Soma, A., Kobayashi, T., Kitabatake, M., Takio, K., Saito, K., Shirouzu, M., Hirao, I., and Yokoyama, S. (2002) Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res. 30*, 4692–9.

3. Wang, W., Takimoto, J. K., Louie, G. V., Baiga, T. J., Noel, J. P., Lee, K.-F., Slesinger, P. a., and Wang, L. (2007) Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nat. Neurosci. 10*, 1063–1072.

4. Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008) Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun. 371*, 818–822.

5. Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science 301*, 964–7.

6. Li, F., Zhang, H., Sun, Y., Pan, Y., Zhou, J., and Wang, J. (2013) Expanding the Genetic Code for Photoclick Chemistry in E. coli , Mammalian Cells, and A. thaliana. *Angew. Chem. Int. Ed. 52*, 9700–9704.

7. Parrish, A. R., She, X., Xiang, Z., Coin, I., Shen, Z., Briggs, S. P., Dillin, A., and Wang, L. (2012) Expanding the Genetic Code of Caenorhabditis elegans Using Bacterial Aminoacyl-tRNA Synthetase/tRNA Pairs. *ACS Chem. Biol. 7*, 1292–1302.

8. Bianco, A., Townsley, F. M., Greiss, S., Lang, K., and Chin, J. W. (2012) Expanding the genetic code of Drosophila melanogaster. *Nat. Chem. Biol. 8*, 748–750.

9. Ernst, R. J., Krogager, T. P., Maywood, E. S., Zanchi, R., Beránek, V., Elliott, T. S., Barry, N. P., Hastings, M. H., and Chin, J. W. (2016) Genetic code expansion in the mouse brain. *Nat. Chem. Biol. 12*, 776–778.

10. Han, S., Yang, A., Lee, S., Lee, H.-W., Park, C. B., and Park, H.-S. (2017) Expanding the genetic code of Mus musculus. *Nat. Commun. 8*, 14568.

11. Liu, C. C., and Schultz, P. G. (2010) Adding New Chemistries to the Genetic Code. *Annu. Rev. Biochem. 79*, 413–444.

12. Chin, J. W. (2014) Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annu. Rev. Biochem. 83*, 379–408.

13. Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) Designing logical codon reassignment – Expanding the chemistry in biology. *Chem. Sci. 6*, 50–69.

14. Hino, N., Hayashi, A., Sakamoto, K., and Yokoyama, S. (2007) Site-specific incorporation of non-natural amino acids into proteins in mammalian cells with an expanded genetic code. *Nat. Protoc. 1*, 2957–2962.

15. Liu, W., Brock, A., Chen, S., Chen, S., and Schultz, P. G. (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat. Methods 4*, 239–244.

16. Srinivasan, G., James, C. M., and Krzycki, J. A. (2002) Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science 296*, 1459–62.

17. Hao, B., Gong, W., Ferguson, T. K., James, C. M., Krzycki, J. A., and Chan, M. K. (2002) A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science 296*, 1462–6.

18. Wan, W., Tharp, J. M., and Liu, W. R. (2014) Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1844*, 1059–1070.

19. Gautier, A., Nguyen, D. P., Lusic, H., An, W., Deiters, A., and Chin, J. W. (2010) Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. *J. Am. Chem. Soc. 132*, 4086–4088.

20. Takimoto, J. K., Adams, K. L., Xiang, Z., and Wang, L. (2009) Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells. *Mol. Biosyst. 5*, 931.

21. Chen, P. R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B. H., and Schultz, P. G. (2009) A Facile System for Encoding Unnatural Amino Acids in Mammalian Cells. *Angew. Chem. Int. Ed. 48*, 4052–4055.

22. Coin, I., Perrin, M. H., Vale, W. W., and Wang, L. (2011) Photo-Cross-Linkers Incorporated into G-Protein-Coupled Receptors in Mammalian Cells: A Ligand Comparison. *Angew. Chem. Int. Ed. 50*, 8077–8081.

23. Shen, B., Xiang, Z., Miller, B., Louie, G., Wang, W., Noel, J. P., Gage, F. H., and Wang, L. (2011) Genetically Encoding Unnatural Amino Acids in Neural Stem Cells

and Optically Reporting Voltage-Sensitive Domain Changes in Differentiated Neurons. *Stem Cells 29*, 1231–1240.

24. Chatterjee, A., Xiao, H., Bollong, M., Ai, H.-W., and Schultz, P. G. (2013) Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells. *Proceedings of the National Academy of Sciences 110*, 11803–11808.

25. Xiao, H., Chatterjee, A., Choi, S.-H., Bajjuri, K. M., Sinha, S. C., and Schultz, P. G. (2013) Genetic Incorporation of Multiple Unnatural Amino Acids into Proteins in Mammalian Cells. *Angew. Chem. Int. Ed. 52*, 14080–14083.

26. Elsässer, S. J., Ernst, R. J., Walker, O. S., and Chin, J. W. (2016) Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat. Methods 13*, 158–164.

27. Schmied, W. H., Elsässer, S. J., Uttamapinant, C., and Chin, J. W. (2014) Efficient Multisite Unnatural Amino Acid Incorporation in Mammalian Cells via Optimized Pyrrolysyl tRNA Synthetase/tRNA Expression and Engineered eRF1. *J. Am. Chem. Soc. 136*, 15577–15583.

28. Zheng, Y., Lewis, T. L., Igo, P., Polleux, F., and Chatterjee, A. (2017) Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues. *ACS Synth. Biol. 6*, 13–18.

29. Swanson, R., Hoben, P., Sumner-Smith, M., Uemura, H., Watson, L., and Soll, D. (1988) Accuracy of in vivo aminoacylation requires proper balance of tRNA and aminoacyl-tRNA synthetase. *Science 242*, 1548–1551.

30. Cohen, S., and Arbely, E. (2016) Single-Plasmid-Based System for Efficient Noncanonical Amino Acid Mutagenesis in Cultured Mammalian Cells. *Chembiochem 17*, 1008–1011.

31. Farr, A., and Roman, A. (1992) A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic Acids Res. 20*, 920–920.

32. Ma, Z.-L., Werner, M., Körber, C., Joshi, I., Hamad, M., Wahle, P., and Hollmann, M. (2007) Quantitative analysis of cotransfection efficiencies in studies of ionotropic glutamate receptor complexes. *J. Neurosci. Res. 85*, 99–115.

33. Chatterjee, A., Sun, S. B., Furman, J. L., Xiao, H., and Schultz, P. G. (2013) A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli. *Biochemistry (Mosc.) 52*, 1828–1837.

34. Lang, K., and Chin, J. W. (2014) Bioorthogonal Reactions for Labeling Proteins. *ACS Chem. Biol. 9*, 16–20.

35. Lang, K., and Chin, J. W. (2014) Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev. 114*, 4764–4806.

36. Plass, T., Milles, S., Koehler, C., Szymański, J., Mueller, R., Wießler, M., Schultz, C., and Lemke, E. a. (2012) Amino Acids for Diels-Alder Reactions in Living Cells. *Angew. Chem. Int. Ed. 51*, 4166–4170.

37. Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D. J., Blackman, M. L., Fox, J. M., and Chin, J. W. (2012) Genetic Encoding of Bicyclononynes and trans -Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic Diels–Alder Reactions. *J. Am. Chem. Soc. 134*, 10317–10320.

38. Nikić, I., Plass, T., Schraidt, O., Szymański, J., Briggs, J. A. G., Schultz, C., and Lemke, E. A. (2014) Minimal Tags for Rapid Dual-Color Live-Cell Labeling and Super-Resolution Microscopy. *Angew. Chem. Int. Ed. 53*, 2245–2249.

39. Peng, T., and Hang, H. C. (2016) Site-Specific Bioorthogonal Labeling for Fluorescence Imaging of Intracellular Proteins in Living Cells. *J. Am. Chem. Soc. 138*, 14423–14433.

40. Lang, K., Davis, L., Torres-Kolbus, J., Chou, C., Deiters, A., and Chin, J. W. (2012)

Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem. 4*, 298–304.

41. Uttamapinant, C., Howe, J. D., Lang, K., Beránek, V., Davis, L., Mahesh, M., Barry, N. P., and Chin, J. W. (2015) Genetic Code Expansion Enables Live-Cell and Super-Resolution Imaging of Site-Specifically Labeled Cellular Proteins. *J. Am. Chem. Soc. 137*, 4602–4605.

42. Seitchik, J. L., Peeler, J. C., Taylor, M. T., Blackman, M. L., Rhoads, T. W., Cooley, R. B., Refakis, C., Fox, J. M., and Mehl, R. a. (2012) Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans -Cyclooctenes. *J. Am. Chem. Soc. 134*, 2898–2901.

43. Nikić, I., Estrada Girona, G., Kang, J. H., Paci, G., Mikhaleva, S., Koehler, C., Shymanska, N. V., Ventura Santos, C., Spitz, D., and Lemke, E. A. (2016) Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy. *Angew. Chem. Int. Ed.* 1–6.

44. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems. *Nat. Chem. Biol. 1*, 13–21.

45. Blackman, M. L., Royzen, M., and Fox, J. M. (2008) Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity. *J. Am. Chem. Soc. 130*, 13518–13519.