Reducing Pyrrolysine tRNA Copy Number Improves Live Cell Imaging of Bioorthogonally Labeled Proteins

Noa Aloush[a], Tomer Schvartz[b], Andres I. König[a], Sarit Cohen[a], Benjamin Tam[1], Barak Akabayov[1], Dikla Nachmias[b], Oshrit Ben-David[a], Natalie Elia[b], Eyal Arbely[a,b,*]

[a] *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*

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

Abstract

Genetic code expansion technology enables the incorporation of non-canonical amino acids (ncAAs) into proteins expressed in live cells. The ncAA is usually encoded by an in-frame stop codon (e.g., TAG) and the methodology relies on the use of an orthogonal aminoacyl tRNA synthetase and its cognate amber suppressor tRNA; for example, the pyrrolysine synthetase/$\mathrm{tRNA}^{\mathrm{Pyl}}_{\mathrm{CUA}}$ (PylT) pair. In such systems, suppression of the in-frame stop codon by the suppressor tRNA is highly dependent on the intracellular concentration of the tRNA. Therefore, multiple copies of *pylT* genes are usually encoded in order to improve ncAA incorporation and protein expression level. However, certain applications of genetic code expansion technology in mammalian cells can benefit from the use of minimal, less invasive, expression systems. For example, live-cell imaging applications, where aminoacylated and labeled suppressor tRNA contributes to high background fluorescence. Therefore, we studied the effect of PylT on live-cell fluorescence imaging of bioorthogonally-labeled intracellular proteins. We found that in COS7 cells, a decrease in *pylT* copy number has no measurable effect on protein expression level and cellular concentration of available PylT. Importantly, we found that reducing *pylT* copy number improves live-cell imaging by enhancing signal-to-noise ratio and reducing immobile PylT population. This enabled us to significantly improve live cell imaging of bioorthogonally labeled intracellular proteins, as well as to co-label two proteins in a cell. Our results indicate that the number of encoded *pylT* genes should be minimized according to the transfected cell line, incorporated ncAA, and the application it is used for.

Keywords: Bioorthogonal labeling, amber suppressor tRNA, super resolution microscopy, inverse electron demand Diels-Alder reaction, tetrazine

1. Introduction

Genetic code expansion technology enables the site-specific incorporation of dozens of non-canonical amino acids (ncAAs) into proteins expressed in live organisms [1–10]. Current methodologies are usually based on the use of 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 studies of ncAA incorporation into proteins expressed in cultured mammalian cells utilized orthogonal aminoacyl-tRNA synthetase/tRNA pairs of bacterial origin, such as *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 archeal pyrrolysyl tRNA synthetase (Pyl-RS) and

*Corresponding author

Email address: arbely@bgu.ac.il (Eyal Arbely)

its cognate amber suppressor $tRNA^{Pyl}_{CUA}$ (PylT) [16, 17] became one of the frequently used orthogonal pairs in cultured mammalian cells [4, 18].

Over the years, methods for expanding the genetic code of cultured mammalian cells were optimized and calibrated extensively [2–4, 19–29]. However, 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 were not identical, making it difficult to compare between different studies (Table S1). That said, these studies significantly improved ncAA incorporation and protein expression levels in mammalian cells. In particular, it was found that the intracellular concentration of suppressor tRNA is a limiting factor in stop codon suppression efficiency and as a result, in overall protein expression level. Moreover, it was demonstrated that high levels of prokaryotic 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 'tRNA cassettes' composed of U6 and/or H1 promoter followed by a suppressor tRNA are encoded in tandem and/or on different plasmids [26–28]. In addition, intracellular level of foreign tRNA such as PylT can be elevated by stabilizing the tRNA; for example, by introducing the U25C and other mutations [24, 29, 30]. These studies suggest that it is crucial for the host system to be able to process the orthogonal tRNA and maintain high intracellular levels of functional tRNA.

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 stop codon suppression efficiency [22, 31]. However, it is difficult to control and measure intracellular concentrations of a tRNA synthetase and its cognate tRNA pair, that are exogenously expressed in transiently transfected cultured mammalian cells. For example,

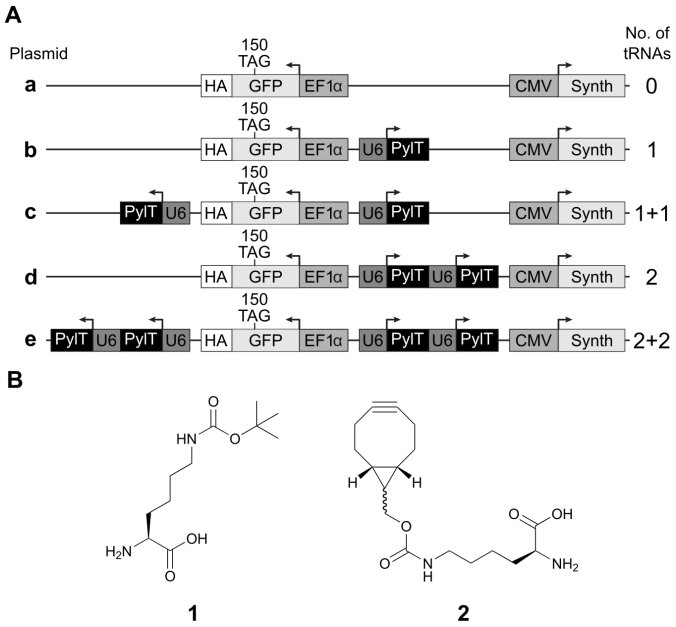

Fig. 1: (A) Schematic representation of plasmids carrying different number of encoded *pylT* genes and $EGFP^{150TAG}$-*HA*. Plasmid **a**, with no encoded tRNA was used as control and plasmid **b** carries 1 copy of *pylT*. Plasmids **c** and **d** carry two encoded copies of *pylT* cloned either separately (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 (marked '2+2'). 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**).

using a viral transfection system it was suggested that efficient amber suppression requires high ratio between suppressor tRNA and its cognate synthetase (using up to 20 copies of encoded suppressor tRNA) [28]. There are also examples for cell lines stably expressing the required genetic components, that were created using the PiggyBac transposon system and two plasmids, each carry 4 copies of *pylT* genes [26]. While these methods offer several advantages, genetic code expansion in transiently transfected cells—were it is more difficult to fine-tune the intracellular levels of synthetases and their cognate tRNAs—is still a frequently used experimental approach.

One of the exciting applications of genetic code expansion technology is the site specific incorporation of ncAAs carrying a bioorthogonal functional group for subsequent chemoselective la-

beling with fluorescent organic dyes, within living cells [32–41]. Such labeling offers a superior alternative to the commonly used fluorescent proteins in fluorescent imaging of live cells [33, 42]. In recent years, bioorthogonal 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 [35–37, 39, 43].

One of the challenging applications in this field, is the live-cell imaging of bioorthogonally labeled intracellular proteins using a *cell-permeable* fluorophore. Such applications suffer from high background fluorescence observed upon addition of the cell-permeable tetrazine-conjugated fluorophores [39, 41]. One of the sources for this background fluorescence (apart from the possible incorporation of the ncAA in response to endogenous amber codons) is the reaction between tetrazine-conjugated fluorophores and aminoacylated tRNAs that leads to the accumulation of fluorophore-labeled tRNAs, mainly in the nucleus [39, 41]. While small, cell permeable fluorophores can be readily washed out of cells, labeled tRNAs are not cell-permeable and may remain in live cells for hours (in contrast to fixed cells with permeabilized membrane).

Several studies in transiently transfected cells found positive correlation between the number of encoded suppressor tRNA genes, intracellular concentration of tRNA, and amber suppression efficiency [4, 22, 27, 44]. However, live-cell fluorescence imaging of bioorthogonally labeled intracellular proteins can benefit from reduction in background fluorescence from aminoacylated tRNAs labeled with organic dyes, as well as minimal effect on cell physiology. Therefore, we decided to study the effect of number of encoded *pylT* genes (and cellular level of PylT) on various parameters of live-cell fluorescence imaging of intracellular proteins. We compared between expression plasmids carrying different copy-number of *pylT* genes, with the aim of increasing signal-to-noise ratio and total number of live cells with site-specifically labeled intracellular proteins.

2. Results

2.1. Protein expression level in COS7 cells is not affected by the number of encoded pylT genes

We have recently demonstrated the high efficiency of a single-plasmid–based system for the incorporation of ncAAs into proteins expressed in transiently transfected cultured mammalian cells [45]. In contrast to transient co-transfection with two or more plasmids, where different populations of cells may take up different combinations of plasmid DNA [46, 47], encoding all the genetic components on one plasmid allows for better control over the ratio between different genes introduced into 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**, Fig. 1A). These plasmids include 0 to 4 copies of the U25C mutant of *Methanosarcina mazei* (*M. mazei*) *pylT* gene controlled by constitutive U6 promoter [30]. 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**, Fig. 1B) as well as with an evolved synthetase (BCN-RS) for the incorporation of bicyclo[6.1.0]nonyne-L-lysine (BCN-Lys, **2**, Fig. 1B). Expression of the synthetase gene is controlled by the human cytomegalovirus (CMV) immediate-early promoter.

With the above described plasmids in hand, we quantified the incorporation level of ncAAs **1** and **2** into 150TAG-mutant of EGFP-HA expressed in COS7 cells (Fig. 2A). We found that reducing the number of encoded *pylT* genes from 4 to 1 had insignificant effect on the expression level of 150TAG-mutant of EGFP-HA in COS7 cells (Fig. 2A). This result was independent of the encoded ncAA, as it was observed when cells were transiently transfected in the presence of ncAA **1** or **2**, with plasmids carrying *Pyl-RS* or *BCN-RS*, respectively.

We therefore decided to focus on plasmids that carry the minimal and maximal number of *pylT* genes (1×PylT plasmid **b** and 2+2×PylT plas-

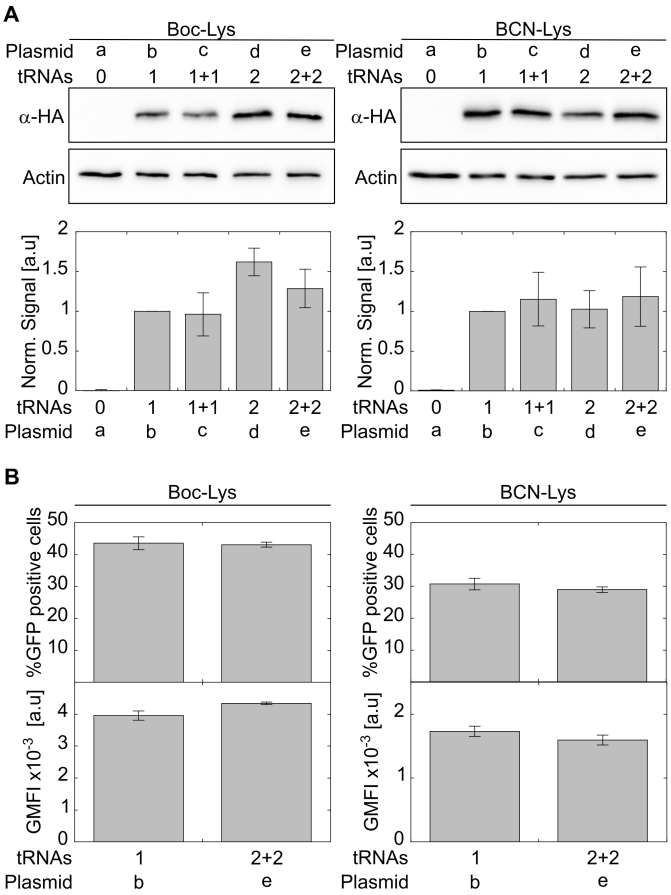

Fig. 2: (A) COS7 cells were transfected with plasmids **a**–**e** carrying either Pyl-RS for the incorporation of **1** or evolved BCN-RS for the incorporation of **2**. Expression levels of $EGFP^{150TAG}$-*HA* were quantified by Western blot using the C-terminus HA-tag, and normalized to actin. Expression levels as function of indicated tRNA copy-number are displayed relative to plasmid carrying 1×tRNA ± SSD (n=5). (B) COS7 cells were transiently transfected with plasmids carrying either Pyl-RS or BCN-RS and 1 or 2+2 copies of tRNA. Transfected cells were 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).

mid **e**, respectively) and quantified the expression level of 150TAG-mutant of EGFP-HA by flow cytometry (Fig. 2B). In excellent agreement with Western blot analyses, single-cell analysis of EGFP expression showed that reducing the number of encoded *pylT* genes from 4 to 1 has no effect on the percentage of viable COS7 cells expressing full length EGFP-HA as well as protein expression level. The data in Fig. 2B was collected 48 h post transfection and similar results were measured 24 h post transfection (Fig. S1). In addition, we found that the reduction in number of encoded *pylT* genes had no effect on the intracellular concentration of PylT in COS7 cells (Fig. S2). Although similar results were recently found in HEK293 cells [29], one would expect positive correlation between the number of encoded *pylT* genes and intracellular concentration of PylT. One possible explanation for the lack of such correlation in COS7 cells, is that the plasmids (mainly the 2+2×PylT plasmid) are not stable in cells. Therefore, we recovered the plasmids 48 hr post-transfection and compared their size by agarose gel electrophoresis (Fig. S3). We found no truncations in the recovered plasmids, indicating that plasmid instability cannot account for the miscorrelation between *pylT* copy number and PylT concentration. Hence, encoding multiple copies of *pylT* genes does not improve the expression level of full-length EGFP with site-specifically incorporated ncAA in transiently transfected COS7 cells.

2.2. Reducing the number of encoded pylT genes increases the number of specifically labeled live cells

In light of the above conclusion, 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, using a cell permeable dye (which is one of the challenging applications in the field). We designed four plasmids to allow for expression and visualization of fluorescently labeled intracellular proteins in live cultured mammalian cells as a function of encoded *pylT* genes (plasmids **f**–**i**, Fig. 3A). As a model for membrane protein labeling, we cloned $EGFP^{150TAG}$-*CAAX* for the expression of EGFP-CAAX anchored to the inner side of plasma membrane *via* prenylation of the C-terminus CAAX motif. In addition, we cloned *HA-α-tubulin*^{45TAG} as a model for cytoskeleton protein labeling, with the aim of labeling cellular microtubules [48]. As a fluorophore for bioorthogonal labeling we

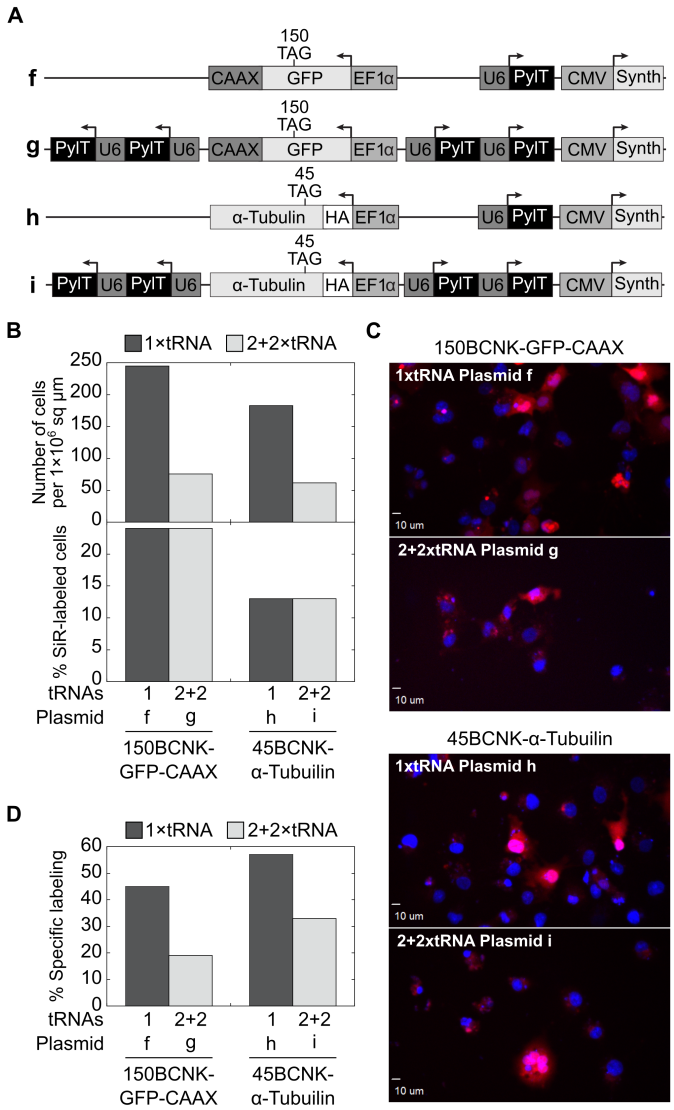

Fig. 3: (A) Schematic representation of 1×PylT or 2+2×PylT plasmids, for the expression of EGFP-CAAX (plasmids **f**, **g**) or α-tubulin (plasmids **h**, **i**) with site-specifically incorporated ncAA **2**. (B) Statistics of live-cell imaging of COS7 cells transfected with plasmids **f**–**i**, stained with Hoechst and labeled with SiR-Tet. Total number of cells found within an area of 1×10^6 sq µm (top) and percent of SiR-labeled cells (bottom) are displayed according to the transfected plasmid. (C) Representative images of live COS7 cells, transfected and labeled as described in B. (D) SiR-labeled cells (n>30) were visualized at ×40 magnification and the cellular location of SiR-labeling was monitored. SiR-positive cells expressing EGFP-CAAX or α-tubulin with clear membrane labeling or labeled microtubules, respectively, were marked as cells with protein-specific labeling. Percent of cells displaying specific labeling are displayed according to the transfected plasmid.

chose the cell-permeable tetrazine-conjugated silicon rhodamine (SiR-Tet) fluorophore (Fig. S4).

COS7 cells were transfected in the presence of ncAA **2** with 1×PylT plasmid **f** or **h**, or with 2+2×PylT plasmid **g** or **i**. Nucleoli were stained with Hoechst, and cells were labeled with SiR-Tet before visualized by live confocal microscopy. Within an area of 1×10^6 sq µm, the total number of cells was highly dependent on the number of encoded *pylT* genes, with ∼3-fold increase in total number of Hoechst-positive cells upon reducing the number of encoded *pylT* genes from 4 to 1 (Fig. 3B top graph, and C). In agreement with flow cytometry data (Fig. 2B), the number of encoded *pylT* genes had no effect on the percent of SiR-positive cells (Fig. 3B bottom graph, and C). That said, since there were ∼3-times more cells following transfection with 1×PylT plasmids, the absolute number of SiR-positive cells was higher in samples transfected with 1×PylT plasmids. Moreover, within a given population of SiR-labeled cells expressing $EGFP^{150TAG}$-*CAAX*, the labeling was membrane-specific in 45% of the cells transfected with 1×PylT plasmid **f**, compared to only 19% of the cells transfected with 2+2×PylT plasmid **g** (Fig. 3D). Similar results were obtained in cells expressing *HA-α-tubulin*^{45TAG} (57% *vs.* 33%). Hence, the use of lower number of encoded *pylT* genes was accompanied by an increase in total number of live cells with protein-specific SiR-labeling.

2.3. The use of minimal number of encoded pylT genes improves the performance of live cell imaging

We assumed that the above described improvement in specific labeling (Fig. 3D) was the result of an increase in signal-to-noise ratio, following the reduction in *pylT* copy number. To test our assumption, we checked how *pylT* copy number affects image quality of live-cell fluorescence imaging of bioorthogonally labeled proteins. When cells were transfected with 2+2×PylT plasmid **g**, specific SiR-labeling of membrane-anchored 150BCNK-EGFP-CAAX, that is notably higher than background, was rare (Fig. 4A, SiR channel). In contrast, using 1×PylT plasmid **f**,

SiR-labeling of membrane-anchored 150BCNK-EGFP-CAAX was visible. To evaluate the extent of background fluorescence, we measured EGFP-CAAX and SiR fluorescence intensities across the plasma membrane (yellow line, Fig. 4A, left panel). When 150BCNK-EGFP-CAAX was expressed from 2+2×PylT plasmid **g**, a sharp peak in intensity level was measured in the GFP channel, representing membrane-anchored 150BCNK-EGFP-CAAX, but practically no notable increase in intensity level was obtained in SiR channel. Instead, SiR fluorescence intensity levels measured in the cytosol were similar or even higher than fluorescence intensity levels measured in the membrane (Fig. 4B). 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 SiR-labeled and membrane-anchored 150BCNK-EGFP-CAAX (Fig. 4B). Moreover, under these conditions, SiR fluorescence intensity levels were significantly higher in the membrane than in the cytosol.

The use of 1 copy of encoded *pylT* gene, instead of 4, also improved image quality of SiR-labeled 45BCNK-α-tubulin expressed in COS7 cells (Fig. 4C). In cells transfected with 2+2×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 45BCNK-α-tubulin, correlates with an increase of 0.45 and 0.26 in signal-to-noise ratio, respectively (Fig. 4D). Compared to the signal-to-noise ratio of GFP-labeled α-tubulin visualized in live COS7 cells (1.75, Fig. 4D, green dashed line), the improvement in signal-to-noise ratio presented here is substantial.

2.4. Reduction in pylT copy-number correlates with reduction in background fluorescence and population of immobile PylT

The SiR-labeling found in the cytosol—and more significantly in the nucleus (Fig. 4A and C)—was ncAA-dependent (Fig. S5) and is probably the result of a reaction between ncAA **2** and SiR-Tet. 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 [39, 41]. In light of the increase in signal-to-noise following the use of 1 *pylT* gene instead of 4 (Fig. 4D), we decided to characterize the effects of encoded *pylT* copy number 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**, Fig. 5A). 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 a co-expressed and labeled protein. Live-cell confocal microscopy revealed that the averaged non-protein-specific fluorescence intensity in the nucleus and cytosol was reduced by ∼25% when *pylT* copy number was reduced from 4 to 1 (Fig. 5B). Hence, reducing the number of encoded *pylT* genes improved the signal-to-noise ratio, probably by reducing background SiR-fluorescence. Furthermore, the equal decrease in nucleus and cytosol confirms that aminoacylated PylT is not entrapped in the nucleus, but rather at a state of dynamic equilibrium between the nucleus and cytosol. The equal reduction of aminoacylated PylT population in the cytosol and the nucleus was measured in cells transfected with plasmid **j** or **k**, suggesting that the equilibrium between the nucleus and cytosol is independent of the number of encoded *pylT* genes.

The reduction in background florescence following the use of one *pylT* gene instead of four correlates with the observed increase in signal-to-noise ratio (Fig. 4D). However, according to Northern blot analysis, the level of PylT in COS7 cells was not affected by the number of encoded *pylT* genes (Fig. S2). In line with these two (seemingly opposing) observations, we noticed that in many cells transfected with a 2+2×PylT plasmid and labeled with SiR-Tet, there are dense 'fluorescent dots' (Fig. S6, white arrows), that are usually absent from cells transfected with a

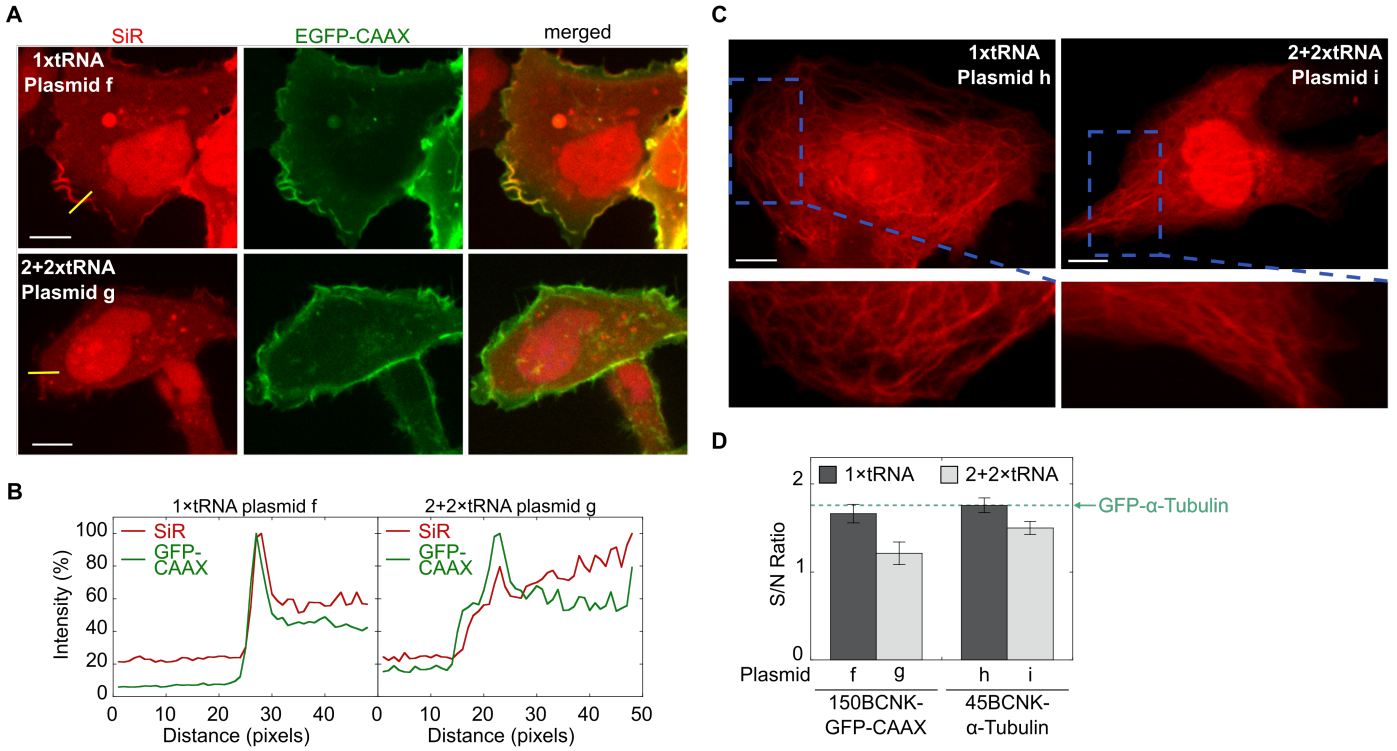

Fig. 4: (A) Live cell imaging of SiR-labeled 150BCNK-EGFP-CAAX expressed in COS7 cells. Left panel: red SiR; center panel: green EGFP; right panel: merge. Scale bar 10 µm. (B) Line intensity profiles of 488 and 640 nm channels (EGFP and SiR, respectively) are presented as percent of maximum value and plotted as a function of distance along a line (yellow line in left panel A). (C) Live cell imaging of SiR-labeled 45BCNK-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. (D) Signal-to-noise ratios calculated from live-cell imaging of SiR-labeled 150BCNK-EGFP-CAAX and 45BCNK-α-tubulin expressed in COS7 cells using plasmids carrying 1 or 4 copies of *pylT* (± SEM). Green dashed line marks the signal-to-noise ratio of GFP-labeled α-tubulin visualized in live COS7 cells under identical conditions.

1×PylT plasmid. Therefore, we hypothesized that the population of PylT in the cell is non-homogeneous, especially when the tRNA is transcribed at high level. To test our hypothesis, we measured the mobility of PylT in COS7 cells using fluorescence loss in photobleaching (FLIP), with the aim of identifying subpopulations of PylT entrapped in large complexes. We compared between SiR-Tet-labeled COS7 cells transfected with 1×PylT plasmid **j** and cells transfected with 2+2×PylT plasmid **k** (Fig. 5C). Over the course of two minutes, for a given number of encoded *pylT* genes, the time-dependent 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 (Fig. 5B). However, overall photobleaching efficiency was dependent on the number of encoded *pylT* genes. 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 (Fig. 5C, left panel). In contrast, when the same FLIP protocol was applied to cells transfected with 2+2×PylT plasmid **k**, fluorescence in the bleached area was reduced to only 40% and fluorescence in the cytosol and nucleus was reduced to about 70% (Fig. 5C, right panel). These results suggest that in COS7 cells transiently transfected with 2+2×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

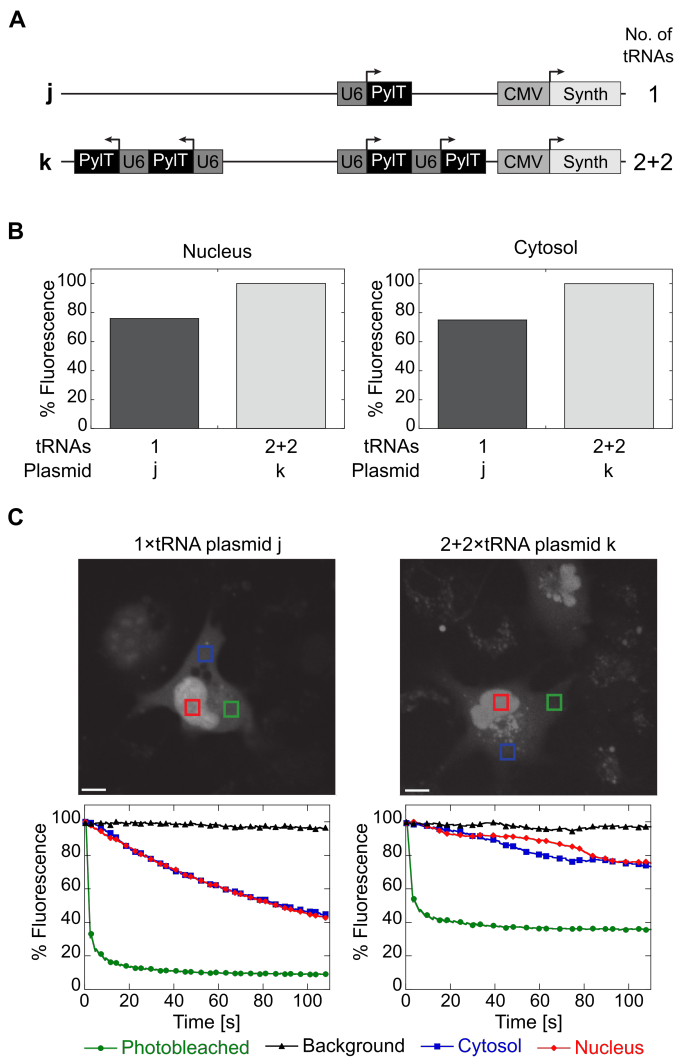

Fig. 5: (A) Schematic representation of plasmids with 1 or 2+2 copies of *pylT* genes and BCN-RS, used in the characterization of non-protein-specific labeling and measurement of aminoacylated PylT dynamics. (B) Relative average SiR-fluorescence measured in the nucleus (left) or cytosol (right) of COS7 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 a function of time. Scale bar 10 µm.

origin and character are yet to be determined.

2.5. *The use of one copy of pylT gene is sufficient for demanding imaging applications*

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

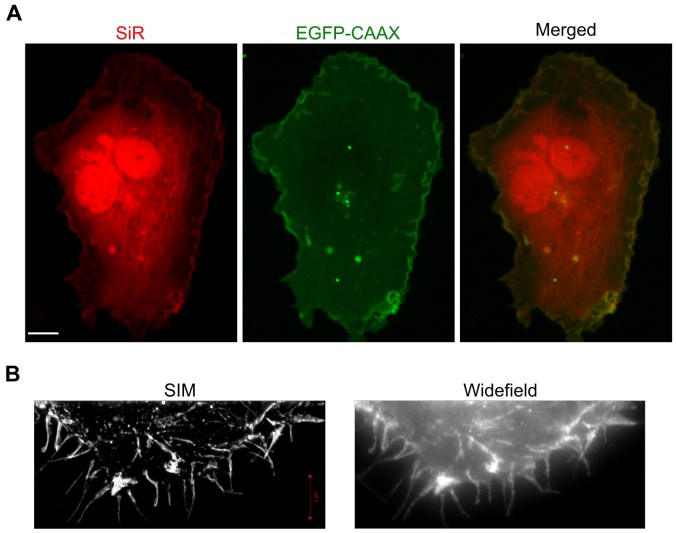

Fig. 6: (A) COS7 cells co-transfected with plasmids **f** and **h** in the presence of BCN-Lys and labeled with SiR-Tet. Using one copy of encoded *pylT* gene, enabled the co-expression and SiR labeling of membrane anchored 150BCNK-EGFP-CAAX and 45BCNK-α-tubulin. Scale bar 10 µm. (B) Zoomed-in regions (rotated by 90 degrees) taken from SIM and widefield imaging of SiR-labeled 150BCNK-EGFP-CAAX (Fig. S7).

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 2+2×PylT plasmids **g** *and* **i** (Fig. 4A), for co-expression and labeling of 150BCNK-EGFP-CAAX *and* 45BCNK-α-tubulin. Co-labeling experiments using 2+2×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 45BCNK-α-tubulin in COS7 cells co-transfected with 1×PylT plasmids **f** and **h** (Fig. 6A). 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 intracellular proteins bioorthogonally labeled with organic fluorophores *via* genetic encoding of ncAAs.

As a final testament to our conclusion that for live-cell imaging applications only one copy of *pylT* gene is better 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, Fig. 6B). 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. Under these conditions, we were able to generate a 3D high-resolution image of the labeled membrane in SIM (Fig. 6B and Fig. S7). This result demonstrates that even with high resolution fluorescence imaging applications that require high labeling density, such as SIM, one copy of encoded *pylT* gene is sufficient for efficient protein expression level and subsequent bioorthogonal labeling of intracellular proteins.

3. Discussion

Genetic code expansion is a powerful technology that enables the expression of 'tailor-made' proteins in live organisms. When combined with site-specific and bioorthogonal labeling of proteins, this technology can revolutionize the way we visualize proteins in live cells. That said, live cell imaging applications may have specific requirements that should be addressed. One of the important challenges in the genetic encoding of a ncAA in response to an in-frame stop codon, is the low suppression efficiency and consequently, low protein expression level. This challenge was addressed by several groups that were able to significantly improve our ability to efficiently incorporate ncAAs into proteins expressed in cultured mammalian cells. These successful studies also contributed to the improvement in fluorescence imaging of bioorthogonally labeled proteins. However, when it comes to the study of biologically relevant questions by visualizing intracellular proteins in live cells, *maximal* protein expression level is not the only important parameter. In such applications, signal-to-noise ratio, proper processing and localization of the labeled protein, and minimal interference with cell physiology are equally important parameters that must be considered. Here, we studied the effect of *pylT* copy-number on imaging quality and signal-to-noise ratio, with the aim of improving these parameters—rather than protein expression yield—with minimal interference with cell physiology.

It is well accepted that amber suppression efficiency correlates with the intracellular concentration of PylT. We found that in COS7 cells transiently transfected with single plasmid, increasing *pylT* copy number has negligible effect on cellular concentration of PylT and consequently on amber suppression efficiency and overall protein expression level. While the correlation between *pylT* copy number and cellular levels of PylT has been reported in several publications, it is difficult to directly compare between these studies due to differences in cell lines, number of expression plasmids and promoter types, DNA delivery methods etc (Table S1). That said, Coin et al. have recently found that similar to the data presented here, the use of one U25C *pylT* gene instead of four has marginal and statistically insignificant effect on the concentration of PylT in HEK293 cells [29]. Importantly, they found that intracellular concentration of PylT can be increased by using an engineered—presumably more stable—mutant of *pylT* (PylT*). These results may be related to our finding of a higher population of immobile PylT in COS7 cells transfected with plasmids carrying four copies of *pylT*, suggesting that excess of PylT is sequestered and becomes unavailable for protein synthesis. The identity of this immobile fraction, and the reasons for the inability of wild type PylT to accumulate at high concentration in COS7 cells are still not clear to us and will be addressed in future studies.

It was recently suggested that the fraction of nuclear PylT can be minimized through the addition of a nuclear export signal (NES) to PylRS and its evolved mutants [41]. In addition, Coin et al. found that the effect of the added NES on protein expression level is dependent on the incorporated ncAA, the transcribed tRNA, and experimental conditions [29]. First, the improvement in protein expression level upon addition of NES was more significant when ncAA **1** was incorporated using wild type PylRS, compared to incorporation of Nε-carboxybenzyl lysine or Nε-trans-cyclooctene lysine using evolved synthetases. Second, the addition of a NES to the synthetase was

beneficial when the protein of interest was expressed at sub-optimal conditions (limited ncAA, 1 copy of tRNA cassette, and co-expression with wild PylT). In contrast, at high concentration of ncAA and co-expression with 4 copies of engineered PylT* (i.e., high concentration of aminoacylated PylT*), the addition of NES to the synthetase did not improve protein expression level. It is not clear if the engineered PylT* is not affected by the nuclear export signal because at such high level of aminoacylated PylT, its cellular concentration is no longer a limiting factor, or because it has a different distribution between the nucleus and cytosol so its concentration in the cytosol is already high. For example, in *Saccharomyces cerevisiae* and COS7 cells, mutation of the C11:G24 base pair of tyrosine tRNA and serine tRNA, respectively, to G11:C24 resulted in nuclear retention of the tRNA [49, 50]. In principal, fluorescence imaging based on bioorthogonal labeling of genetically encoded ncAAs should benefit from any reduction in the amount of labeled PylT in the nucleus. However, an increase in cytosolic fraction of PylT may interfere with live-cell imaging using cell-permeable dyes due to labeling of cytosolic and aminoacylated PylT (in contrast to imaging of fixed cells and non cell permeable dyes). It is therefore important to study how a shift in PylT equilibrium between the nucleus and cytosol affects signal-to-noise ratio and cytosolic background fluorescence in live cell imaging applications using cell permeable dyes.

Taken together, we improved image quality and signal-to-noise ratio of live cell imaging of SiR-labeled proteins expressed in transiently transfected COS7 cells, by reducing *pylT* copy number. The reduction in number of encoded *pylT* genes was also accompanied by an increase in total number of cells available for live cell imaging, following transfection and multi-step labeling protocol. Hence, we suggest that in live cell imaging of bioorthogonally labeled proteins using genetic code expansion technology and cell permeable dyes, the number of encoded *pylT* genes should be (and can be) minimized based on the cell line, incorporated ncAA, and labeled protein.

Acknowledgments

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.).

Bibliography

[1] L. Wang, A. Brock, B. Herberich, P. G. Schultz, Expanding the Genetic Code of Escherichia coli, Science 292 (5516) (2001) 498–500. doi:10.1126/science.1060077.

[2] K. Sakamoto, A. Hayashi, A. Sakamoto, D. Kiga, H. Nakayama, A. Soma, T. Kobayashi, M. Kitabatake, K. Takio, K. Saito, M. Shirouzu, I. Hirao, S. Yokoyama, Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells, Nucleic Acids Research 30 (21) (2002) 4692–9. doi:10.1093/nar/gkf589.

[3] W. Wang, J. K. Takimoto, G. V. Louie, T. J. Baiga, J. P. Noel, K.-F. Lee, P. a. Slesinger, L. Wang, Genetically encoding unnatural amino acids for cellular and neuronal studies, Nature Neuroscience 10 (8) (2007) 1063–1072. doi:10.1038/nn1932.

[4] T. Mukai, T. Kobayashi, N. Hino, T. Yanagisawa, K. Sakamoto, S. Yokoyama, Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases, Biochemical and Biophysical Research Communications 371 (4) (2008) 818–822. doi:10.1016/j.bbrc.2008.04.164.

[5] J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang, P. G. Schultz, An expanded eukaryotic genetic code, Science 301 (5635) (2003) 964–7. doi:10.1126/science.1084772.

[6] F. Li, H. Zhang, Y. Sun, Y. Pan, J. Zhou, J. Wang, Expanding the Genetic Code for Photoclick Chemistry in E. coli , Mammalian Cells, and A. thaliana, Angewandte Chemie International Edition 52 (37) (2013) 9700–9704. doi:10.1002/anie.201303477.

[7] A. R. Parrish, X. She, Z. Xiang, I. Coin, Z. Shen, S. P. Briggs, A. Dillin, L. Wang, Expanding the Genetic Code of Caenorhabditis elegans Using Bacterial Aminoacyl-tRNA Synthetase/tRNA Pairs, ACS Chemical Biology 7 (7) (2012) 1292–1302. doi:10.1021/cb200542j.

[8] A. Bianco, F. M. Townsley, S. Greiss, K. Lang, J. W. Chin, Expanding the genetic code of Drosophila melanogaster, Nature Chemical Biology 8 (9) (2012) 748–750. doi:10.1038/nchembio.1043.

[9] R. J. Ernst, T. P. Krogager, E. S. Maywood, R. Zanchi, V. Beránek, T. S. Elliott, N. P. Barry, M. H. Hastings, J. W. Chin, Genetic code expansion in the mouse brain, Nature Chemical Biology 12 (10) (2016) 776–778. doi:10.1038/nchembio.2160.

[10] S. Han, A. Yang, S. Lee, H.-W. Lee, C. B. Park, H.-S. Park, Expanding the genetic code of Mus musculus, Nature Communications 8 (2017) 14568. doi:10.1038/ncomms14568.

[11] C. C. Liu, P. G. Schultz, Adding New Chemistries to the Genetic Code, Annual Review of Biochemistry 79 (1) (2010) 413–444. doi:10.1146/annurev.biochem.052308.105824.

[12] J. W. Chin, Expanding and Reprogramming the Genetic Code of Cells and Animals, Annual Review of Biochemistry 83 (1) (2014) 379–408. doi:10.1146/annurev-biochem-060713-035737.

[13] A. Dumas, L. Lercher, C. D. Spicer, B. G. Davis, Designing logical codon reassignment Expanding the chemistry in biology, Chem. Sci. 6 (1) (2015) 50–69. doi:10.1039/C4SC01534G.

[14] N. Hino, A. Hayashi, K. Sakamoto, S. Yokoyama, Site-specific incorporation of non-natural amino acids into proteins in mammalian cells with an expanded genetic code, Nature Protocols 1 (6) (2007) 2957–2962. doi:10.1038/nprot.2006.424.

[15] W. Liu, A. Brock, S. Chen, S. Chen, P. G. Schultz, Genetic incorporation of unnatural amino acids into proteins in mammalian cells, Nature Methods 4 (3) (2007) 239–244. doi:10.1038/nmeth1016.

[16] G. Srinivasan, C. M. James, J. A. Krzycki, Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA, Science 296 (5572) (2002) 1459–62. doi:10.1126/science.1069588.

[17] B. Hao, W. Gong, T. K. Ferguson, C. M. James, J. A. Krzycki, M. K. Chan, A new UAG-encoded residue in the structure of a methanogen methyltransferase, Science 296 (5572) (2002) 1462–6. doi:10.1126/science.1069556.

[18] W. Wan, J. M. Tharp, W. R. Liu, Pyrrolysyl-tRNA synthetase: An ordinary enzyme but an outstanding genetic code expansion tool, Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1844 (6) (2014) 1059–1070. arXiv:NIHMS150003, doi:10.1016/j.bbapap.2014.03.002.

[19] A. Gautier, D. P. Nguyen, H. Lusic, W. An, A. Deiters, J. W. Chin, Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells, Journal of the American Chemical Society 132 (12) (2010) 4086–4088. doi:10.1021/ja910688s.

[20] J. K. Takimoto, K. L. Adams, Z. Xiang, L. Wang, Improving orthogonal tRNA-synthetase recognition for efficient unnatural amino acid incorporation and application in mammalian cells, Molecular BioSystems 5 (9) (2009) 931. doi:10.1039/b904228h.

[21] P. R. Chen, D. Groff, J. Guo, W. Ou, S. Cellitti, B. H. Geierstanger, P. G. Schultz, A facile system for encoding unnatural amino acids in mammalian cells., Angewandte Chemie (International ed. in English) 48 (22) (2009) 4052–5. doi:10.1002/anie.200900683.

[22] I. Coin, M. H. Perrin, W. W. Vale, L. Wang, Photo-Cross-Linkers Incorporated into G-Protein-Coupled Receptors in Mammalian Cells: A Ligand Comparison, Angewandte Chemie International Edition 50 (35) (2011) 8077–8081. doi:10.1002/anie.201102646.

[23] B. Shen, Z. Xiang, B. Miller, G. Louie, W. Wang, J. P. Noel, F. H. Gage, L. Wang, Genetically Encoding Unnatural Amino Acids in Neural Stem Cells and Optically Reporting Voltage-Sensitive Domain Changes in Differentiated Neurons, Stem Cells 29 (8) (2011) 1231–1240. doi:10.1002/stem.679.

[24] A. Chatterjee, H. Xiao, M. Bollong, H.-W. Ai, P. G. Schultz, Efficient viral delivery system for unnatural amino acid mutagenesis in mammalian cells, Proceedings of the National Academy of Sciences 110 (29) (2013) 11803–11808. doi:10.1073/pnas.1309584110.

[25] H. Xiao, A. Chatterjee, S.-H. Choi, K. M. Bajjuri, S. C. Sinha, P. G. Schultz, Genetic Incorporation of Multiple Unnatural Amino Acids into Proteins in Mammalian Cells, Angewandte Chemie International Edition 52 (52) (2013) 14080–14083. doi:10.1002/anie.201308137.

[26] S. J. Elsässer, R. J. Ernst, O. S. Walker, J. W. Chin, Genetic code expansion in stable cell lines enables encoded chromatin modification, Nature Methods 13 (2) (2016) 158–164. doi:10.1038/nmeth.3701.

[27] W. H. Schmied, S. J. Elsässer, C. Uttamapinant, J. W. Chin, Efficient Multisite Unnatural Amino Acid Incorporation in Mammalian Cells via Optimized Pyrrolysyl tRNA Synthetase/tRNA Expression and Engineered eRF1, Journal of the American Chemical Society 136 (44) (2014) 15577–15583. doi:10.1021/ja5069728.

[28] Y. Zheng, T. L. Lewis, P. Igo, F. Polleux, A. Chatterjee, Virus-Enabled Optimization and Delivery of the Genetic Machinery for Efficient Unnatural Amino Acid Mutagenesis in Mammalian Cells and Tissues, ACS Synthetic Biology 6 (1) (2017) 13–18. doi:10.1021/acssynbio.6b00092.

[29] R. Serfling, C. Lorenz, M. Etzel, G. Schicht, T. Böttke, M. Mörl, I. Coin, Designer tRNAs for efficient incorporation of non-canonical amino acids by the pyrrolysine system in mammalian cells, Nucleic Acids Researchdoi:10.1093/nar/gkx1156.

[30] A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao, P. G. Schultz, A Versatile Platform for Single- and Multiple-Unnatural Amino Acid Mutagenesis in Escherichia coli, Biochemistry 52 (10) (2013) 1828–1837. doi:10.1021/bi4000244.

[31] R. Swanson, P. Hoben, M. Sumner-Smith, H. Uemura, L. Watson, D. Soll, Accuracy of in vivo

aminoacylation requires proper balance of tRNA and aminoacyl-tRNA synthetase, Science 242 (4885) (1988) 1548–1551. doi:10.1126/science.3144042.

[32] K. Lang, J. W. Chin, Bioorthogonal Reactions for Labeling Proteins, ACS Chemical Biology 9 (1) (2014) 16–20. doi:10.1021/cb4009292.

[33] K. Lang, J. W. Chin, Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins, Chemical Reviews 114 (9) (2014) 4764–4806. doi:10.1021/cr400355w.

[34] T. Plass, S. Milles, C. Koehler, J. Szymański, R. Mueller, M. Wießler, C. Schultz, E. a. Lemke, Amino Acids for Diels-Alder Reactions in Living Cells, Angewandte Chemie International Edition 51 (17) (2012) 4166–4170. doi:10.1002/anie.201108231.

[35] K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox, J. W. Chin, Genetic Encoding of Bicyclononynes and trans -Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic DielsAlder Reactions, Journal of the American Chemical Society 134 (25) (2012) 10317–10320. doi:10.1021/ja302832g.

[36] I. Nikić, T. Plass, O. Schraidt, J. Szymański, J. A. G. Briggs, C. Schultz, E. A. Lemke, Minimal Tags for Rapid Dual-Color Live-Cell Labeling and Super-Resolution Microscopy, Angewandte Chemie International Edition 53 (8) (2014) 2245–2249. doi:10.1002/anie.201309847.

[37] T. Peng, H. C. Hang, Site-Specific Bioorthogonal Labeling for Fluorescence Imaging of Intracellular Proteins in Living Cells, Journal of the American Chemical Society 138 (43) (2016) 14423–14433. doi:10.1021/jacs.6b08733.

[38] K. Lang, L. Davis, J. Torres-Kolbus, C. Chou, A. Deiters, J. W. Chin, Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction, Nature Chemistry 4 (4) (2012) 298–304. doi:10.1038/nchem.1250.

[39] C. Uttamapinant, J. D. Howe, K. Lang, V. Beránek, L. Davis, M. Mahesh, N. P. Barry, J. W. Chin, Genetic Code Expansion Enables Live-Cell and Super-Resolution Imaging of Site-Specifically Labeled Cellular Proteins, Journal of the American Chemical Society 137 (14) (2015) 4602–4605. doi:10.1021/ja512838z.

[40] J. L. Seitchik, J. C. Peeler, M. T. Taylor, M. L. Blackman, T. W. Rhoads, R. B. Cooley, C. Refakis, J. M. Fox, R. a. Mehl, Genetically Encoded Tetrazine Amino Acid Directs Rapid Site-Specific in Vivo Bioorthogonal Ligation with trans -Cyclooctenes, Journal of the American Chemical Society 134 (6) (2012) 2898–2901. doi:10.1021/ja2109745.

[41] I. Nikić, G. Estrada Girona, J. H. Kang, G. Paci, S. Mikhaleva, C. Koehler, N. V. Shymanska, C. Ventura Santos, D. Spitz, E. A. Lemke, Debugging Eukaryotic Genetic Code Expansion for Site-Specific Click-PAINT Super-Resolution Microscopy, Angewandte Chemie International Edition (2016) 1–6doi:10.1002/anie.201608284.

[42] J. A. Prescher, C. R. Bertozzi, Chemistry in living systems, Nature Chemical Biology 1 (1) (2005) 13–21. doi:10.1038/nchembio0605-13.

[43] M. L. Blackman, M. Royzen, J. M. Fox, Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand DielsAlder Reactivity, Journal of the American Chemical Society 130 (41) (2008) 13518–13519. doi:10.1021/ja8053805.

[44] M. Buvoli, A. Buvoli, L. a. Leinwand, Suppression of Nonsense Mutations in Cell Culture and Mice by Multimerized Suppressor tRNA Genes, Molecular and Cellular Biology 20 (9) (2000) 3116–3124. doi:10.1128/MCB.20.9.3116-3124.2000.

[45] S. Cohen, E. Arbely, Single-Plasmid-Based System for Efficient Noncanonical Amino Acid Mutagenesis in Cultured Mammalian Cells, ChemBioChem 17 (11) (2016) 1008–1011. doi:10.1002/cbic.201500681.

[46] A. Farr, A. Roman, A pitfall of using a second plasmid to determine transfection efficiency, Nucleic Acids Research 20 (4) (1992) 920–920. doi:10.1093/nar/20.4.920.

[47] Z.-L. Ma, M. Werner, C. Körber, I. Joshi, M. Hamad, P. Wahle, M. Hollmann, Quantitative analysis of co-transfection efficiencies in studies of ionotropic glutamate receptor complexes, Journal of Neuroscience Research 85 (1) (2007) 99–115. doi:10.1002/jnr.21096.

[48] T. Schvartz, N. Aloush, I. Goliand, I. Segal, D. Nachmias, E. Arbely, N. Elia, Direct fluorescent-dye labeling of α-tubulin in mammalian cells for live cell and superresolution imaging, Molecular Biology of the Cell 28 (21) (2017) 2747–2756. doi:10.1091/mbc.E17-03-0161.

[49] J. D. Cleary, D. Mangroo, Nucleotides of the tRNA D-stem that play an important role in nuclear-tRNA export in Saccharomyces cerevisiae, The Biochemical journal 347 Pt 1 (1) (2000) 115–22. doi:10.1042/0264-6021:3470115.

[50] S. C. Chafe, D. Mangroo, Scyl1 Facilitates Nuclear tRNA Export in Mammalian Cells by Acting at the Nuclear Pore Complex, Molecular Biology of the Cell 21 (14) (2010) 2483–2499. doi:10.1091/mbc.E10-03-0176.