Dual stop codon suppression in mammalian cells with genomically integrated genetic code expansion machinery

Birthe Meineke,[1,2]* Johannes Heimgärtner,[1,2] Rozina Caridha,[1,2] Matthias Block,[1] Kyle J Kimler,[1] Maria F Pires,[1] Michael Landreh[3] and Simon J Elsässer[1,2,]*

[1]Science for Life Laboratory, Karolinska Institutet, Department of Medical Biochemistry and Biophysics, Division of Genome Biology, Stockholm 17165, Sweden

[2]Ming Wai Lau Centre for Reparative Medicine, Stockholm node, Karolinska Institutet, Stockholm 17165, Sweden

[3]Department of Microbiology, Tumor and Cell Biology, Science for Life Laboratory, Karolinska Institutet, Stockholm 17165, Sweden

*Correspondence: birthe.meineke@scilifelab.se; simon.elsasser@scilifelab.se

ABSTRACT

Genetic code expansion via stop codon suppression is a powerful strategy to engineer proteins. Suppressor tRNAs are aminoacylated with noncanonical amino acids (ncAAs) by dedicated aminoacyl-tRNA synthetases (aaRS) and direct ncAA incorporation site-specifically during translation. These pairs of tRNA/aaRS must be orthogonal to the host's tRNAs, aaRS and natural amino acids. Pyrrolysyl-tRNA (PylT)/PylRS pairs from methanogenic archaea, as well as engineered tRNA/aaRS pairs derived from bacteria, are used for genetic code expansion in mammalian cells. Amber suppression is routinely achieved by transient introduction of the components leading to short-term and heterogeneous expression. Here, we demonstrate that stable integration of tRNA/aaRS genes allows for efficient, genetically encoded ncAA incorporation in diverse mammalian cell lines. We extend a general plasmid design and PiggyBac (PB) integration strategy developed for the *Methanosarcina mazei* PylT/PylRS pair to genomic integration of two tRNA/aaRS pairs of bacterial origin. We further explore suppression of ochre and opal stop codons and parallel incorporation of two distinct ncAAs, both accessible for click chemistry, by dual suppression in stable cell lines. Clonal selection allows for isolation of cells with high dual suppression efficiency and dual site-specific fluorescent labeling of a cell surface receptor using bioorthogonal click chemistries on live mammalian cells.

INTRODUCTION

The genetic code assigns all possible triplet codons to one of the canonical amino acids (sense codons) or to a translation stop signal (nonsense codons). Expansion of the genetic code, that is genetic encoding of noncanonical amino acids (ncAAs), requires reprogramming of a codon. Most commonly the amber (TAG) stop codon is repurposed, using an engineered tRNA/aminoacyl-tRNA synthetase (aaRS) pair to deliver an ncAA-charged tRNA to the ribosome. The amber codon is the least abundant (23% in the human genome) and most efficiently suppressed stop codon, but the other two stop codons UUA (ochre) and UGA (opal) have also been exploited for ncAA incorporation in mammalian cells and enabled dual ncAA incorporation via dual suppression (Xiao et al. 2013; Zheng et al. 2017) (Meineke et al. 2018, 2020; Serfling et al. 2018; Beránek et al. 2019; Osgood et al. 2022). Quadruplet codon suppression has been proven possible in worms and mammalian cells as well (Niu et al. 2013; Xi et al. 2022). The pyrrolysyl-tRNA (PylT) and PylT synthetase (PylRS) pair from methanogenic archaea, is widely used for amber suppression in mammalian cells. In the Methanosarcinaceae family of archaea, specific amber codons in mRNAs coding for enzymes of the methanogenic pathway are suppressed by PylT to encode pyrrolysine (Pyl) (Rother and Krzycki 2010). PylT and PylRS have characteristics that make them ideal for genetic code expansion in model organisms: the conserved fold of PylT allows specific interaction with PylRS independent of anticodon identity (Srinivasan et al. 2002) (Ambrogelly et al. 2007) and the relatively promiscuous amino acid binding pocket of PylRS readily accommodates structurally diverse ncAAs. This broad substrate specificity has been further increased by protein engineering, over 100 ncAAs can be incorporated by PylRS variants (reviewed in (Wan et al. 2014)). Importantly, the PylT/RS pair is orthogonal to prokaryotic and eukaryotic tRNAs or aaRS (Suzuki et al. 2017). These features have made the pairs from *Methanosarcina mazei* (*Mma*) and *Methanosarcina bakeri* (*Mba*) widely used tools for site-specific integration of a plethora of noncanonical amino acids (ncAAs) in bacterial, eukaryotic, mammalian host systems and animals (Chin 2014).

Another approach to expand the genetic code by amber suppression relies on engineering of canonical tRNA/aaRS pairs from one organism in a host of a different clade of life. This has been achieved for both tyrosyl-tRNA synthetase (TyrRS) and leucyl-tRNA synthetase (LeuRS) from *E. coli* by changing the anticodon of the cognate tRNA (TyrT and LeuT, respectively) to CUA and sequential rounds of selection for orthogonality and efficient amber suppression. An

engineered *Eco*TyrRS mutant, AzFRS, has been shown to effectively incorporate p-azido-phenylalanine (AzF) with either *Eco*TyrT^{CUA} or *B. stearothermophilus* TyrT^{CUA} (*Bst*Tyr^{CUA}) (Sakamoto et al. 2002) (Chin et al. 2003) (Liu et al. 2007). *Eco*LeuRS-derived Anap-2C (AnapRS) paired with *Eco*LeuT^{CUA}, allows introduction of the minimal fluorescent ncAA 3-(6-acetylnaphthalen-2-ylamino)-2-amino-propanoic acid (Anap) in mammalian cells (Lee et al. 2009; Chatterjee et al. 2013).

In mammalian cell culture, components of the amber suppression machinery, i.e. the tRNA/aaRS pair and reporter or gene of interest with amber codon, are commonly delivered by transient transfection. Lipofection methods allow easy delivery of high copy numbers of the expression vectors into the cell, but the resulting populations are heterogeneous in expression levels and endogenous transcripts are outcompeted by the transgene (Muerdter et al. 2018). For prolonged and homogenous protein production, extended expression experiments, selection of homologous expression within a cell population and other applications, amber suppression cell lines with all components of the genetic code expansion machinery integrated in their genome, are desirable. We have used PiggyBac (PB) transgenesis to achieve genetic code expansion from stable integrants previously (Elsässer et al. 2016; Elsässer 2018; van Husen et al. 2021; Lee et al. 2009). PB transposition utilizes a deconstructed transposon cassette: cotransfection of a PB transposase (PBT) expression plasmid with a second plasmid containing an integration cassette flanked by inverted repeats recognised by PBT leads to random genomic integration of the cassette at TTAA sites in the host genome. Integrants are selected by antibiotic resistance markers included in the inserted cassette. While PB transgenesis is known to allow multiple integration events per cell (Ding et al. 2005), expression of the protein of interest is expected to be much lower and closer to high-expressing endogenous genes as compared to a transient transfection experiment. Therefore, stable genetic code expansion is commonly viewed as inefficient. Lower, genomic levels of the gene of interest may increase the competition with off-target suppression at endogenous amber codons, raising the question how selectivity towards the transgene's stop codon can be achieved in an endogenous expression context (Bartoschek et al. 2021). It is thus important to comprehensively characterize stable suppression systems for understanding the opportunities and limitations of transgenic cell lines. In the present study, we aimed to create a panel of representative cell lines and systematically characterize efficiency and selectivity of amber suppression using PB transgenesis. We demonstrate that stable genetic code expansion greatly varies in efficiency by cell line, but can

rival transient expression in specific settings and enable efficient single and dual stop codon suppression.

RESULTS

Integrating amber suppression machinery into the genome of human cell lines by PB transgenesis

We have previously developed a plasmid system for generation of stable amber suppression cell lines (pAS plasmids) using *Mma* PylT/RS (PylT/RS hereafter) (Figure 1A) (Elsässer et al. 2016; Elsässer 2018): One plasmid for PylRS expression and a second encoding a sfGFP^{150TAG} (GFP^{150TAG}) reporter gene, each controlled by the EF1α promoter. Both plasmids carry four tandem repeats of 7SK PolIII-controlled PylT genes. A resistance marker is expressed from an internal ribosome entry site (IRES) in the *PylS* or GFP transcripts to allow selection of integrants. The expression cassette is flanked on both ends by insulator sequences and inverted repeats for PB transposition. This plasmid design allowed for highest suppression efficiency in transient expression compared to other systems (Zhou et al. 2020) and we have successfully employed pAS plasmids for both transient transfection and stable integration in the past (Meineke et al. 2018, 2020; Lafranchi et al. 2020; van Husen et al. 2021; Elsässer et al. 2016).

Common immortalized animal cell lines differ in their characteristics with regard to transgene expression and ease of transfection. We compare PB-mediated integration of our amber suppression system, consisting of *PylT, PylS* and the GFP^{150TAG} reporter, across a panel of commonly used cell lines. The two plasmids were contransfected with the PBT plasmid into human embryonic kidney derived cell lines HEK293 and HEK293T, human colon cancer cell line HCT116, human melanoma A375 and osteosarcoma cell line U-2 OS as well as primate cell line COS-7. The resulting stable polyclonal pools were grown in the presence or absence of N-ε-[(2-methyl-2-cyclopropene-1-yl)-methoxy] carbonyl-L-lysine (CpK) for 24 h. GFP was only expressed in presence of CpK as evident from anti-GFP immunoblotting (Figure 1B, even numbered lanes). We used strain-promoted inverse electron-demand Diels-Alder cycloaddition (SPIEDAC) between CpK and tetrazine-silicon rhodamine (tetSiR) in lysates to visualize CpK incorporation (Elliott et al. 2014). A single band corresponding to GFP$^{150CpK\text{-}tetSiR}$ showed selective ncAA incorporation in all cell lines generated (Figure 1B). Faint additional bands indicated low level incorporation of ncAA into various endogenous proteins. The amount of GFP^{150CpK} produced greatly differed between cell lines. HCT116-derived cells produced the

strongest GFP signal, followed by HEK293T and HEK293 cells. Curiously, PylRS protein levels were increased or decreased upon ncAA addition in some cell lines and PylRS expression levels did not correlate with production of full-length GFP. We did not further investigate this phenomenon. While lysates were normalized by total protein content before gel separation, GAPDH levels varied, presumably because of cell type-dependent expression patterns.

PylT gene copy number is limiting in HEK293T stable amber suppression cell lines

To determine which components are limiting amber suppression, we transfected a relatively inefficient HEK293T amber suppression cell line with different pAS constructs to increase gene copy number of PylT, PylRS or both. A control pAS plasmid did not change GFP^{150CpK} or PylRS expression (lanes 1-3, Figure 1C, D). Transfection with pAS expressing PylRS lead to strongly increased FLAG-PylRS protein level, but no increased amber suppression efficiency, as GFP signal did not increase (lane 4, Figure 1C, D). Addition of pAS with either only 4xPylT or 4xPylT/PylRS increased GFP^{150CpK} expression (lanes 5 and 6, Figure 1C, D). The GFP fluorescence increased 2.5-fold with transient transfection of pAS 4xPylT and tripled with pAS with 4xPylT/PylRS. Amber suppression did not improve by further increasing the PylT gene dosage with pAS 8xPylT, or by using reported optimized PylT variants M15 or A2-1 (Serfling et al. 2018; Jewel et al. 2023). Like wild-type PylRS, chimeric *Mba/Mma*PylRS IPYE mutant had no effect on amber suppression, despite its reported higher catalytic rate (Bryson et al. 2017) (Figure S1A). This indicates that PylT levels can limit amber suppression when they are exclusively expressed from randomly integrated pAS transgenes. The comparably high levels of amber suppression achieved in transient transfection are therefore linked to high PylT gene copy number, while PylRS expression and activity are not rate-limiting.

Comparison of transient and stable amber suppression

We further compared amber suppression efficiency of stable polyclonal HEK293, HEK293T and HCT116 cell lines with their transient transfected counterparts. After 24 h incubation with CpK, transiently transfected HEK293T cells express GFP^{150CpK} very heterogeneously: flow cytometry shows 60% GFP-positive cells, but the level of GFP fluorescence varies greatly within that population. Stable HEK293T cells produced a less extreme but more homogeneous level of GFP reporter fluorescence between 10^2-10^3 arbitrary units (Figure 1E). While HEK293 and HCT116 cells showed much reduced transfection efficiency, stable cell lines exhibit efficient suppression across the entire population. Comparing mean GFP fluorescence per cell

(indicative of the total yield of GFP), stable HCT116 cells produced almost as much GFP^{150CpK} as transiently transfected HEK293T cells, suggesting that transgenic HCT116 are an excellent expression hosts for producing ncAA-containing proteins in mammalian cells (Figure S1B).

We further confirmed this finding using a different ncAA and PylRS variant. The chimeric *Mma/Mba* PylRS variant AbKRS-chIPYE (Bryson et al. 2017; Ai et al. 2011; Gautier et al. 2010), engineered for incorporation of the photocrosslinking ncAA 3′-azibutyl-N-carbamoyl-lysine (AbK), also produced an efficiently amber-suppressing polyclonal population of HCT116 cells (Figure S1C, D).

PB transgenesis with non-PylT/RS amber suppression pairs in HCT116 cells

We generated pAS constructs for AzF and Anap incorporation analogous to our PylT/RS two-plasmid system (Figure S2A). AzFRS was combined with *Bst*TyrT^{CUA} (TyrT^{CUA}/AzFRS) and AnapRS was combined with *Eco*LeuT^{CUA} (LeuT^{CUA}/AnapRS), both pAS were designed with a geneticin (neomycin) selectable marker. The cognate tRNA/GFP^{150TAG} reporters were also generated.

We compared amber suppression efficiency of transient transfection and PB transgenesis for AzFRS-mediated amber suppression in HCT116 cells, measuring GFP fluorescence by flow cytometry after 24 h with AzF. Transient transfection of HCT116 cells was again inefficient. Selection with geneticin and blasticidin for PB mediated integration of TyrT^{CUA}/AzFRS and TyrT^{CUA}/GFP^{150TAG} produced a heterogeneous pool, but the fraction of GFP positive cells in presence of AzF increased from 4.7% in transient transfection to 39.2% (Figure 2A). Fluorescence-assisted cell sorting (FACS) of the polyclonal pool for single, strongly GFP-expressing cells, isolated clonal populations with very high amber suppression efficiency (Figure S2B). Evidently, the TyrT^{CUA}/AzFRS pair can support extremely efficient amber suppression from integrated transgenes.

We similarly compared the polyclonal pool of stable LeuT^{CUA}/AnapRS and LeuT^{CUA}/GFP^{150CpK} integrants with transiently transfected cells. Here, we assessed both GFP and Anap fluorescence by flow cytometry (Figure 2B). The polyclonal pool of stable cell lines showed strong GFP and Anap fluorescence when incubated with the ncAA for 24 h, indicating expression of GFP^{150Anap} (Figure 2B, S2C). A fluorigenic ncAA like Anap ideally allows direct detection of the protein of interest, however unincorporated and misincorporated ncAA can contribute substantially to signal intensity. To assess the contribution of misincorporated Anap in endogenous proteins, we generated a HCT116 cell line with integrated LeuT^{CUA}/AnapRS without

amber suppression reporter. When cultured in presence of Anap for 24 h this AnapRS-only integrant cell line produced no GFP fluorescence, but Anap fluorescence was clearly detectable by flow cytometry despite extended ncAA-washout before measurement (Figure S2C-E). Evidently, Anap was retained in the cell in aminoacyl-tRNA complexes or incorporated at off-target amber codons even in the absence of a dedicated target protein.

Total GFP yield as determined by western blot was marked higher in stable AzFRS and AnapRS amber suppression HCT116 cell lines (Figure 2C). We confirmed AzFRS expression via a C-terminal FLAG tag after stable integration, while AnapRS, expressed in its originally described tagless form ((Lee et al. 2009; Chatterjee et al. 2013)), could not be probed.

Efficient amber-, ochre- and opal-suppressing clonal populations generated by PB transgenesis in HCT116 cells

The amber codon is most widely used in genetic code expansion in mammalian cells, because it is the least abundant stop codon and amber suppression efficiency is superior to suppression of either of the other two stop codons, ochre (UAA) or opal (UAG) in transient transfections (Zheng et al. 2017; Meineke et al. 2018, 2020). We sought to explore if genomic integration would allow for selection and isolation of efficient ochre and opal suppressor populations, a prerequisite for incorporating more than one ncAA site-specifically in response to two different stop codons in a stable dual suppression cell line. We generated 4x7SK PylT cassettes for expression of ochre and opal suppressor PylTs, $PylT^{UUA}$ and $PylT^{UCA}$, and combined them with PylRS and their cognate GFP reporter GFP^{150TAA} or GFP^{150TGA}, respectively. In transient transfection, neither ochre nor opal suppression were efficient in HCT116 cells (Figure 3A, left column). Selection for stable transfectants (with the same antibiotic concentration for all three stop codon constructs) yielded polyclonal pools of HCT116 cell lines that express GFP^{150CpK}. GFP yield of the ochre and opal suppression polyclonal populations was however more than an order of magnitude lower (mean fluorescence 81 and 109 arbitrary units, respectively) than amber suppression (mean at 1687 arbitrary units) (Figure 3A, center column; Figure S3A). We isolated clones from the polyclonal population by flow cytometry, selecting for single cells with strong GFP signal after 24 h incubation with CpK. Six amber-suppressing clones, nine ochre- and one opal-suppressing clones were recovered and compared for GFP expression by flow cytometry (Figure S3A, B). For each clone, addition of CpK resulted in homogenous GFP fluorescence at higher intensity than observed for the polyclonal pool. The most efficiently suppressing clones for each stop codon are shown in Figure 3A (right column). Despite the clonal selection, maximal ochre and

opal suppression remained roughly one order of magnitude lower than amber suppression (Figure S3B). This confirmed the observation from transient transcription experiments that ochre and opal codons are more difficult to suppress.

Efficient suppression of the target stop codon is typically considered the most important variable for optimization, but the selectivity for ncAA labeling applications also crucially depends on low misincorporation at endogenous stop codons (Bartoschek et al. 2021). In order to compare the on- versus off-target incorporation of ncAA, we compared the best clones for each stop codon after extended CpK incorporation. Cell lysates were labeled by SPIEDAC with tetSiR during lysis (Figure 3B). The most prominent band corresponds to $GFP^{150CpK\text{-}tetSiR}$ for each of the three stop codon-suppressing cell lines. The ochre suppression clonal cell lysate also featured a distinct band, suggesting a particularly strongly labeled endogenous protein. It is interesting to note, that endogenous opal codons do not appear to be more efficiently suppressed by PylT than the other two stop codons, despite the opal codon being known to show highest leakiness, i.e. readthrough with endogenous tRNAs (Floquet et al. 2012; Schmied et al. 2014; Palomar-Siles et al. 2022). In summary, productive ochre and opal suppression is possible to achieve in stable settings. Yet, as expected, amber suppression of the reporter GFP is markedly more efficient while exhibiting similarly low levels of incorporation into endogenous proteins.

Dual suppression clonal cell line for amber and ochre co-suppression

In the past years, several approaches have successfully implemented dual suppression of two different stop codons in transiently transfected mammalian cells, allowing for site-specific incorporation of two distinct ncAAs (Zheng et al. 2017; Xiao et al. 2013; Meineke et al. 2018, 2020; Beránek et al. 2019; Serfling et al. 2018). Encouraged by selection of ochre and opal suppressor cell lines expressing high levels of GFP^{150CpK}, we decided to investigate if dual suppression can be implemented in a stable setting. The compatibility of TyrT/RS and LeuT/RS pairs from *E. coli* in combinations with archaeal PylT/PylRS has been explored in detail using transient transfection (Zheng et al. 2017). Choosing the two most efficiently suppressed stop codons (Figure 3A), we explored the combination of the ochre suppressor $PylT^{UUA}$/PylRS pair with a compatible orthogonal amber suppressor tRNA/aaRS pair for dual integration. We integrated $PylT^{UUA}$/RS, $TyrT^{CUA}$/AzFRS and a $GFP^{102TAG150TAA}$ dual suppression reporter into HCT116 cells (Figure 4A). After co-selection with puromycin, geneticin and blasticidin, a small subpopulation of the polyclonal pool showed increased GFP fluorescence after 48 h with CpK and AzF. We isolated single cells with high GFP fluorescence by FACS and recovered clonal

populations with a varied range of dual suppression efficiency (Figure S4A). The most efficient clone showed robust, CpK- and AzF-dependent expression of $GFP^{102AzF150CpK}$ (Figure 4B, C and S4B, C). No cross-incorporation of ncAA was observed as judged from the lack of GFP-positive cells after incubation with only one of the ncAAs (Figure 4B, C). We purified GFP expressed in presence of AzF and CpK from the clonal cell line (yield 80 ng/ml culture) and confirmed incorporation of both ncAAs by intact mass spectrometry (Figure 4D). In addition to CpK, $GFP^{102AzF150CpK}$ contains AzF, which can also be used for bioorthogonal labeling. The azido group can react with alkynes in Cu(I) catalysed azide-alkyne cycloaddition (CuAAC) or with strained alkynes in copper independent strain promoted azide-alkyne cycloaddition (SPAAC). We visualized purified $GFP^{102AzF150CpK}$ by SPIEDAC and SPAAC with tetSiR and dibenzylcyclooctyne-PEG4-5/6-tetramethyl-rhodamine (DBCO-TAMRA), respectively (Figure 4E). In summary, stable dual suppression cell lines are useful for protein production.

Amber suppression cell lines for cell surface receptor expression

We have so far described PB cell line generation with an easily detectable, well expressed and stable GFP reporter. To demonstrate the robustness of our approach, we chose a range of other proteins involved in different cellular processes. We generated stable amber suppression cells lines in HCT116 cells for expression of a number of intracellular proteins of interest, either with premature TAG or STELLA-N-ter tag and C-terminal HA-tag (Lafranchi et al. 2020): The cytoskeletal protein tubulin, heterochromatin associated DAXX, the histone variant CENP-A and the microproteins PIGBOS and SARS2-CoV-M (Figure S5A). We used the PylRS “AF” variant in these cell lines to incorporate axial trans-cyclooct-2-ene-L-lysine (TCO*K) (Nikić et al. 2014; Yanagisawa et al. 2008). Polyclonal stable populations were selected and all proteins were detectable by immunoblotting for their HA-tag, with expected variation in expression levels. SPIEDAC labeling with tetSiR or tetrazine-AF488 (tet488) in cell lysates shows specific bands for the TCO*K-labeled proteins according to their respective size, albeit endogenous proteins appear to be labeled with similar or even higher intensity (Figure S5B, C). Fluorescence microscopy of fixed and permeabilized cells labeled with methyl-tetrazine-BDP-FL (metetBDPFL) and counterstained for HA further showed that bioorthogonal labeling created a strong ambient background signal across cytosol and nucleus while a specific signal from the amber suppressed protein was only marginally brighter (Figure S5D). This is most evident for the Golgi-restricted CoV-M protein (Lafranchi et al. 2020). Therefore, the strong expression and selective suppression of our GFP reporter must be considered an outlier compared to many

biologically relevant proteins that are desirable to label using amber suppression. Evidently, the known large plasmid copy number, ten to hundred thousands copies per cell (Fliedl et al. 2015), after transient transfection provides a much larger window for selective suppression of the target amber codon than expression from stable PB integrants. Nevertheless, we considered that mislabeling of endogenous C-terminal amber stop codons may generate predominantly intracellular background and using a cell-impermeable dye may allow selective labeling of target proteins on the cell surface.

To test this hypothesis, we used two cell surface receptors for generation of stable amber suppression HCT116 cell lines: First, the class B GPCR corticotropin-releasing factor type 1 receptor (CRFR1). CRFR1 A95TAG has previously been used in amber suppression studies for ncAA mediated bioorthogonal labeling on the surface of live cells (Coin et al. 2013; Serfling et al. 2018; Meineke et al. 2020, 2021). Second, a minimal synthetic Notch receptor (SynNotch). An amber codon at N1703 (in the *Notch1* gene, position N442 in our construct) places the ncAA on the cell surface close to the membrane spanning helix of the C-terminal fragment generated by proteolysis during maturation of the receptor (Meineke et al. 2020).

Polyclonal stable integrants of each receptor were cultured with TCO*K for 48 h and live cells were SPIEDAC-labeled with tetrazine-Cy3 (tetCy3) and subsequently fixed for microscopy. CRFR1^{95TCO*K}- and SynNotch^{442TCO*K}-expressing cells showed a strong, TCO*K-dependent tetCy3 signal on the surface of cells (Figure 5A, B). Expression levels were much more homogeneous across the population compared to transient transfection (Meineke et al. 2020). The same CRFR1^{95TAG} and SynNotch^{442TAG} cell lines were labeled live with tet488 and lysates were labeled with tetSiR for comparison (Figure 5C). While the tet488 signal was highly specific for the cell surface receptor and overlapped with the HA-immunostaining signal, tetSiR in lysates preferentially labeled endogenous proteins. In summary, stable amber suppression cell lines provide limited opportunity for selective bioorthogonal labeling of intracellular proteins, but enable efficient suppression and labeling of extracellular receptors, offering an ideal platform for biophysical studies of membrane receptors in their native environment.

Dual site-specific ncAA incorporation and dual color labeling by combined amber and ochre suppression in transgenic HCT116 cells

Building on our promising results for combining amber and ochre suppression with PylT^{UUA}/RS and TyrT^{CUA}/AzFRS for the GFP reporter, we attempted generation of stable dual suppression cell lines in HCT116 cells for SynNotch. The SynNotch^{204TAA442TAG} amber and ochre double

mutant can incorporate two distinct ncAAs for dual bioorthogonal derivatization when transiently cotransfected with appropriate, orthogonal tRNA/aaRS pairs. The incorporation sites are designed to place the two ncAAs in close spatial proximity to each other, but on separate parts of the mature, proteolytically processed SynNotch receptor (Meineke et al. 2020).

Initial attempts to generated PB transgenic cell lines for SynNotch^{204TAA442TAG}-HA failed, even when we exchanged the ochre suppressor PylT for the M15 mutant (Serfling et al. 2018), which we had found to enhance ochre suppression efficiency before (Meineke et al. 2020). We recovered PylRS and AzFRS expressing populations after selection, but were unable to detect ncAAs or HA-tag (data not shown). To facilitate selection of integrants with SynNotch receptor expression by fluorescence-activated cell sorting (FACS), we exchanged the C-terminal HA-tag for a GFP-tag. We transfected HCT116 cells with M15^{UUA}/PylRS, TyrT^{CUA}/AzFRS, M15^{UUA}/SynNotch^{204TAA442TAG}-GFP and PBT. After selection with geneticin, puromycin and blasticidin, GFP fluorescence could be detected in a small number of cells after culturing in presence of AzF and CpK for 3 days. From this population, we isolated single cells by FACS for highest GFP fluorescence (Figure S6A). One clonal population showed increased GFP fluorescence with both ncAAs over a 12 to 72 h time course (Figure S6B). Flowcytomery and imaging of this clone confirmed that GFP fluorescence was not the result of cross-incorporation, readthrough with endogenous amino acids or a secondary translation start, as it localized to the cell membrane and depended on addition of both CpK and AzF (Figure 6A, S6C).

We validated the functionality of the two ncAAs for dual bioorthogonal labeling with tetCy3 and DBCO-AF647 (DBCO647) by SPIEDAC and SPAAC, respectively, on SynNotch^{204CpK442AzF} -GFP purified from the clonal population (Figure 6B). In SDS-PAGE the two proteolytic fragments of SynNotch are separated. The Cy3 and 647 fluorescence signals are isolated which demonstrates that each receptor domain is selectively labeled and confirms site-specific incorporation of the two ncAAs. As expected, the DBCO647 signal from the AzF labeled C-terminal transmembrane domain overlapped with anti-GFP immunostaining signal. We observe two diffuse bands for the glycosylated N-terminal domain, due to partial proteolytic cleavage at the TEV site (Meineke et al. 2020). A small proportion of unprocessed SynNotch^{204CpK442AzF}-GFP was also captured, visible as a faint high molecular weight band in both fluorescence channels and immunoblotting. We validated assignment of the bands by transient transfection of SynNotch^{204TAA442TAG}-GFP with wild type PylT/RS and SPIEDAC labeling on live cells, which resulted in the same band pattern (Figure S6D).

Live-cell dual color bioorthogonal labeling of a receptor in stable cell lines.

Next, we wondered if SynNotch^{204CpK442AzF}-GFP, produced by dual suppression in our stable clonal cell line, could also be labeled on live cells. Cells were grown in presence of both ncAAs for 72 h, labeled with tetCy3 and alkyne-AF647 (alk647) and imaged for fluorescence of both dyes (Figure 6C). We detected overlapping, ncAA dependent cell surface signals for all three fluorophores associated with SynNotch^{204CpK442AzF}-GFP. The Cy3, 647 and GFP fluorescences were dependent on addition of both ncAA. When CuAAC was performed on live cells, both GFP and Cy3 signals were decreased; highlighting the potential of copper to damage cellular proteins even at low concentration (Figure S6E). Using a copper-free SPAAC reaction with DBCO-467 as as alternative for labeling AzF, however, was not successful, as it leads to strong unspecific labeling of endogenous proteins (Figure S6F) (van Geel et al. 2012) .

Clonal SynNotch^{204CpK442AzF}-GFP expressing cells labeled by SPIEDAC and CuAAC under live conditions yield the same band pattern observed for affinity purified SynNotch^{204CpK442AzF}-GFP when separated by SDS-PAGE (Figure 6D). Two ncAA-dependent bands with tet488-fluorescence are visible for the N-terminal SynNotch domain, due to partial proteolysis at the TEV-site. An additional higher molecular weight band represented an artifact of tet488 surface labeling since it was present in all dye-treated samples, even in absence of ncAAs (Figure 6D). After CuAAC labeling a single fluorescent band is visible, corresponding to C-terminal TMD-GFP, comprising N442AzF specifically labeled by alk647 (Figure 6D). SPAAC labeling of SynNotch^{204CpK442AzF}-GFP with DBCO647 lead to a diffuse signal across the entire lane, owing to DBCO reactivity with protein residues other than the incorporated azide ncAA (Figure S6G). Taken together our results demonstrate that dual suppression and bioorthogonal derivatization of two distinct ncAAs with two mutually orthogonal reactions is possible in a stable cell line with integrated suppression machinery. The resulting fluorescence signal strength was limited by the low efficiency of dual stop codon suppression and by the limited choice of mutual orthogonal reactions.

DISCUSSION

Systematic evaluation of suppression systems and mammalian host cell lines enabled us to generate stable cell lines with expanded genetic codes for suppression of all three stop codons using PB transgenesis. We established a universal two-plasmid design for generating mammalian cell lines with the choice of three distinct orthogonal tRNA/Synthetase pairs. This versatility should enable incorporation of virtually any genetically encoded ncAA described to

date in stable cell lines. While stop codon suppression, especially for ochre and opal codons, can be initially low after drug selection, highly efficient clones can be isolated. In the present study. We did not systematically assess DNA copy number and resulting relative expression levels of PylRS, PylT and the target protein in a larger number of clones. Consequently, a universal recipe for the optimal dosage of tRNA, aaRS and gene of interest cannot be derived from our data, and the best condition may be highly context specific with respect to cell line, target protein, and ncAA used. Nevertheless, we find that maximizing tRNA and gene of interest expression are determinants of ncAA incorporation efficiency and selectivity. In addition, position and context of the stop codon have been shown to greatly influence suppression efficiency and hence are important parameters to optimize (Bartoschek et al. 2021).

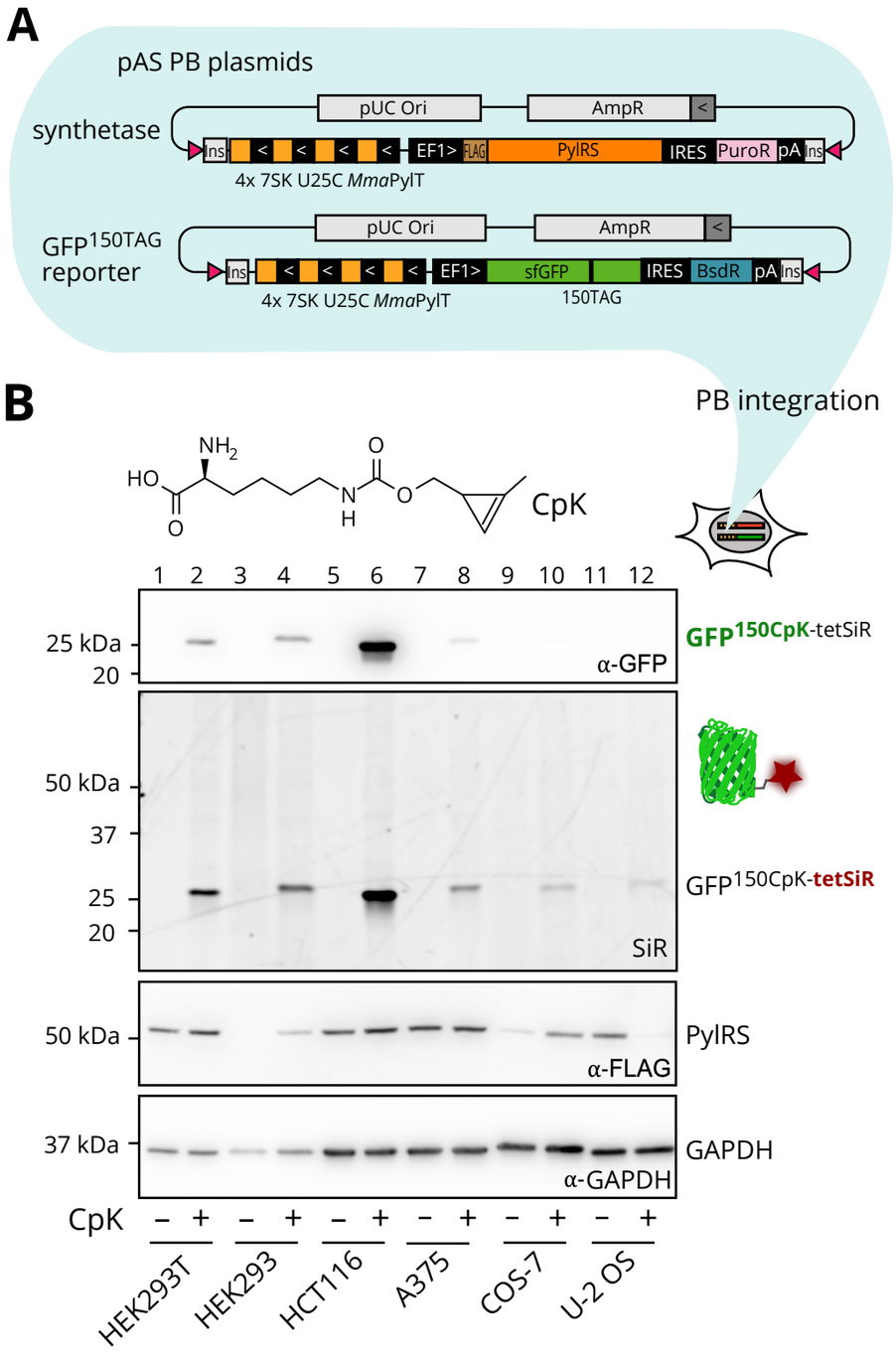
Our approach benefits from the efficiency of PB transgenesis. Optimal clones can be isolated by selection without a priori knowledge of the ideal copy number configuration. As an alternative to genomic integration, an episomal self-replicating plasmid has been used to long-term express the amber suppression machinery in human hematopoietic stem cells (Shao et al. 2020). This approach, however, will likely require continuous selection to maintain stable transgene copy number and efficient amber suppression. Stable integration and clonal selection, including by PB transposition, has been also used to generate amber suppression CHO cell lines efficiently expressing ncAA-modified antibodies for drug conjugation (Oller-Salvia et al. 2018; Axup et al. 2012).

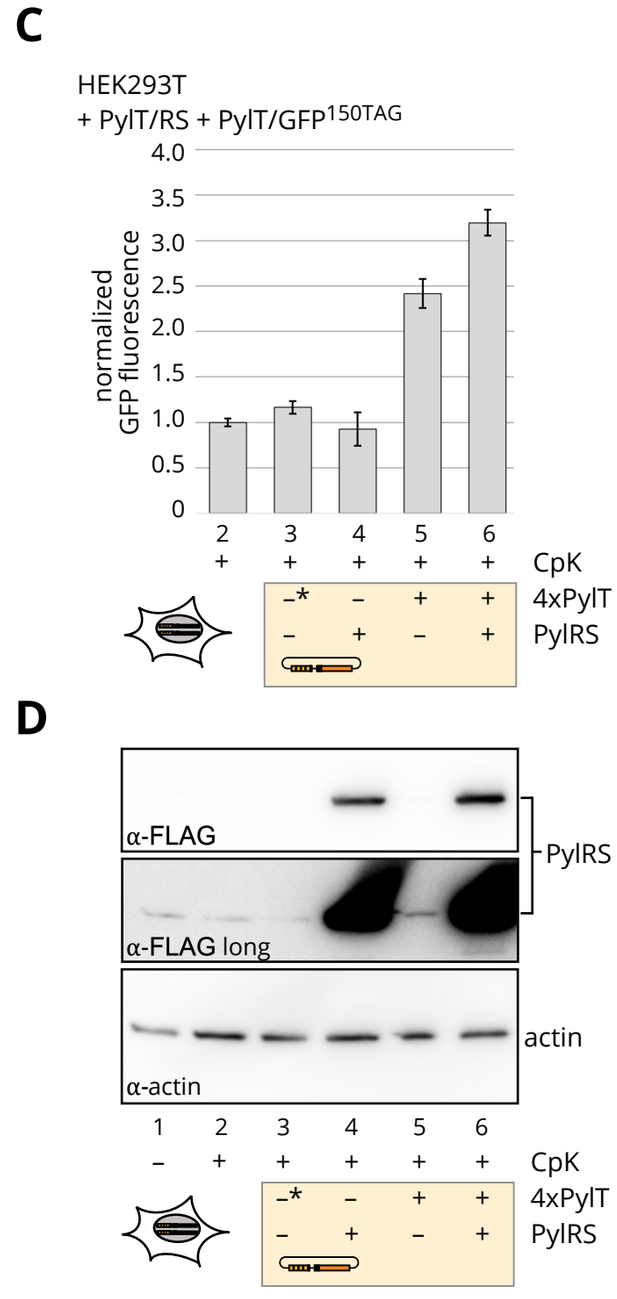
Building on the toolbox created in this study, we proceeded to demonstrate site-specific incorporation and chemical labeling of two distinct ncAAs in a stable mammalian cell line. Our study highlights opportunities and limitations of dual suppression systems, requiring two selective, efficient and mutual orthogonal incorporation machineries. The relatively lower efficiency of ochre and opal stop codon suppression limits overall yield of dually suppressed protein (Zheng et al. 2017; Meineke et al. 2018, 2020; Osgood et al. 2022). We chose AzF for encoding the CuAAC handle because of its efficient incorporation (Figure 2A, S2B), but this likely limited the fluorescent labeling efficiency on live cells. We have previously shown that a genetically encoded picolyl-azide provide superior CuAAC reactivity with azide dyes (Meineke et al. 2021), but orthogonal machineries to introduce picolyl-azide and TCO*K or CpK do not exist to date.

It is an important goal of the field to achieve selective labeling of amber codons in target proteins expressed at endogenous level or even endogenously edited genes. Our experiments with the GFP reporter demonstrate that, under most optimal conditions, endogenously

expressed proteins can be ncAA labeled with high efficiency and good selectivity. However, our data also underscores that such an ideal scenario is unlikely to apply to the average cellular protein. Our methodology provides most attractive applications for labeling proteins on the plasma membrane. Site-specifically incorporated ncAAs on the cell surface face little competition from misincorporated ncAAs at endogenous stop codons due to the fact that C-termini of plasma membrane proteins are mostly intracellular or posttranslationally processed. Our stable genetic code expansion technology, together with fluorescent bioorthogonal labeling on the cell surface, provides a valuable toolbox to investigate the localization, movement and clustering of membrane receptors on the cell surface. Dual-color labeling will enable new applications for studying conformational dynamics and multimeric assemblies of cell surface proteins.

Figure 1


A
pAS PB plasmids
synthetase
pUC Ori
AmpR
EF1>
FLAG
PylRS
IRES
PuroR
pA
4x 7SK U25C MmaPylT
GFP150TAG reporter
sfGFP
BsdR
150TAG
PB integration
B
CpK
GFP150CpK-tetSiR
GFP150CpK-tetSiR
PylRS
GAPDH
25 kDa
20
50 kDa
37
25
20
50 kDa
37 kDa
α-GFP
SiR
α-FLAG
α-GAPDH
CpK
HEK293T
HEK293
HCT116
A375
COS-7
U-2 OS


C
HEK293T
+ PylT/RS + PylT/GFP150TAG
normalized GFP fluorescence
CpK
4xPylT
PylRS
D
α-FLAG
α-FLAG long
α-actin
PylRS
actin
CpK
4xPylT
PylRS

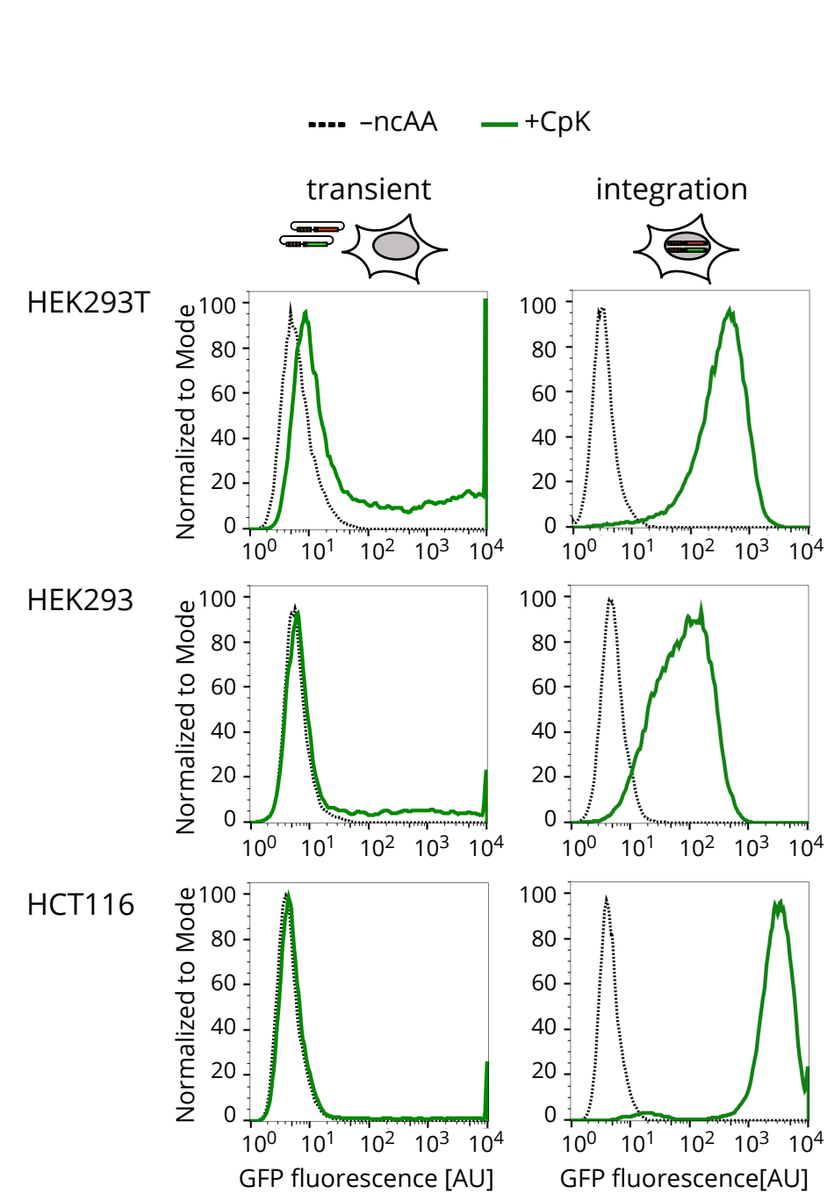

E
–ncAA
+CpK
transient
integration
HEK293T
HEK293
HCT116
Normalized to Mode
GFP fluorescence [AU]

Figure 1 - PB transgenesis integrates PylT/RS amber suppression machinery into the host cell genome.

A) Schematic representation of pAS (Amber Suppression) PB transgenesis plasmids. N-terminally FLAG-tagged PylRS (PylRS) is expressed under control of the EF1α promoter. An internal ribosome entry site (IRES), followed by a puromycin resistance selectable marker (PuroR), is positioned immediately downstream of PylRS. A cassette of four tandem repeats of PylT each controlled by h7SK PolIII promoter are positioned upstream of the EF1α promoter. The entire PylT/RS cassette is flanked by insulator sequences (Ins) and 3′ and 5′ inverted repeats for genomic integration by PB transposase (pink triangle). The GFP^{150TAG} amber suppression reporter pAS construct has the same architecture, but contains a blasticidin resistance marker (Bsd) instead.

B) Immunoblot comparing GFP^{150CpK} reporter expression in different mammalian PB integrant cell lines. The chemical structure of CpK is shown at the top of the panel. HEK293T (lanes 1 and 2), HEK293 (lanes 3 and 4), HCT116 (lanes 5 and 6), A375 (lanes 7 and 8), COS-7 (lanes 9 and 10) and U-2 OS (lanes 11 and 12) PB amber suppression integrants were grown with 0.5 mM CpK (+, even numbered lanes) or without ncAA (–, odd numbered lanes) for 24 h. Soluble lysates were SPIEDAC labeled with 1 μM tetSiR and normalized for total protein content. SiR signal was visualized at 630 nm in-gel after SDS-PAGE. Immunostaining for GFP expression levels, FLAG-PylRS and GAPDH.

C) Quantification of GFP fluorescence in lysates of a low selected (2 μg/ml puromycin, 500 μg/ml blasticidin), stable PylT/RS and GFP^{150TAG} integrant, HEK293T cell line. Where indicated (yellow shading), the cell line was transiently transfected with different pAS plasmids bearing 4xPylT and/or PylRS variants. Transfections were performed in quadruplicate with 0.2 mM CpK addition for 24 h in triplicate. GFP fluorescence signal is normalized to untransfected cells cultured in presence of 0.2 mM CpK for 24 h, background fluorescence (–CpK) values were subtracted. Error bars indicate standard deviation. Plasmids transfected by lanes: 2 no transfection, lane 3: control transfection with a pAS control plasmid, lane 4: transfection with pAS (4xPylT), lane 5: transfection with pAS_*Mma*PylRS (PylRS), lane 6: transfection with pAS_4x*Mma*PylT/RS (4xPylT and 4xPylRS).

D) Immunoblot for PylRS in lysates shown in C). Lanes 1 and 2: no transfection without (–) and with (+) CpK, others lanes as in C). Soluble lysates were separated by SDS-PAGE. Immunostaining for FLAG-PylRS and β-actin.

E) Flow cytometry comparing amber suppression in HEK293T (top), HEK293 (middle) and HCT116 (bottom) after transient transfection (left) and PB transgenesis (right). Cells were transfected transiently with PylT/RS and the PylT/GFP^{150TAG} reporter. Alternatively, the PylT/RS and the PylT/GFP^{150TAG} reporter were integrated by PB transgenesis using the same plasmids. GFP fluorescence from the sfGFP^{150TAG} reporter was measured after 24 h without ncAA (dashed black line) or in presence of 0.2 mM CpK (green line).

Figure 2

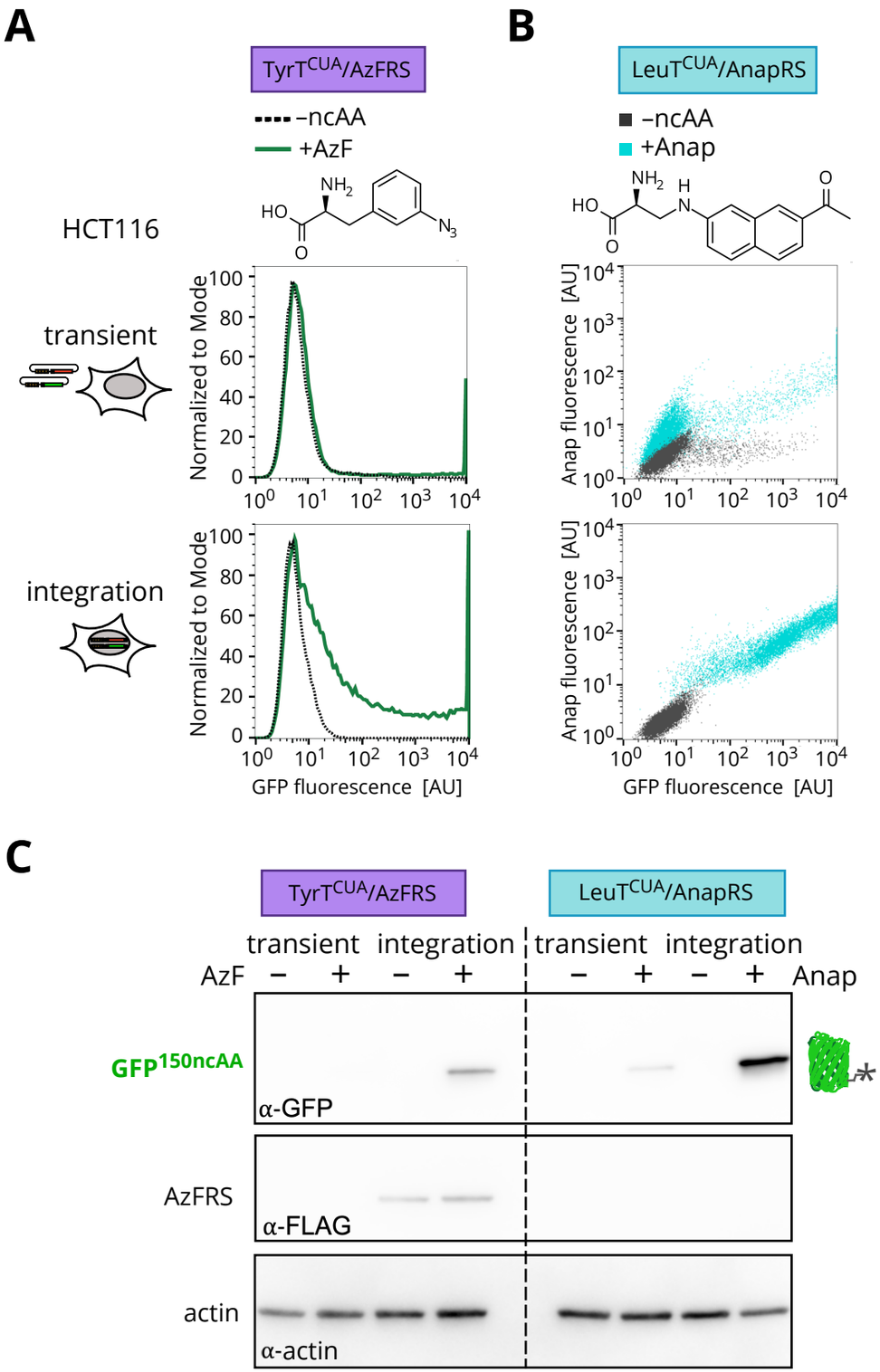

A
B
TyrT^CUA/AzFRS
–ncAA
+AzF
LeuT^CUA/AnapRS
–ncAA
+Anap
HCT116
transient
integration
Normalized to Mode
Anap fluorescence [AU]
GFP fluorescence [AU]
C
TyrT^CUA/AzFRS
LeuT^CUA/AnapRS
transient
integration
AzF
Anap
GFP^150ncAA
α-GFP
AzFRS
α-FLAG
actin
α-actin

Figure 2 - PB integration of $TyrT^{CUA}$/AzFRS and $LeuT^{CUA}$/AnapRS in HCT116 cells.

A) Flow cytometry of amber suppression cell line generated with *E. coli* TyrRS-derived AzFRS. The chemical structure of AzF is shown. HCT116 cells were transiently transfected with $TyrT^{CUA}$/AzFRS and corresponding $TyrT^{CUA}$/GFP^{150TAG} reporter (transient, top). The same plasmids were integrated by PB transgenesis (integration, bottom). GFP fluorescence from the $sfGFP^{150TAG}$ reporter was measured after 24 h without ncAA (dashed black line) or in presence of 0.5 mM AzF (green line, chemical structures shown).

B) Flow cytometry of amber suppression cell line generated with *E. coli* LeuRS-derived AnapRS. The chemical structure of Anap is shown. HCT116 cells were transiently transfected with $LeuT^{CUA}$/AnapRS and cognate $LeuT^{CUA}$/GFP^{150TAG} reporter (transient, top). The same plasmids were integrated by PB transgenesis (integration, bottom). GFP and Anap fluorescence were measured without ncAA (black) or after 24 h incubation with 0.05 mM Anap and a 2 h chase with ncAA-free medium (cyan) before measurement.

C) Immunoblotting of lysates prepared from transiently transfected and PB integrant HCT116 cells cultured in presence of ncAA as in A) and B). Samples were separated by SDS-PAGE and immunostained for $GFP^{150ncAA}$, FLAG-AzFRS and an actin loading control.

Figure 3

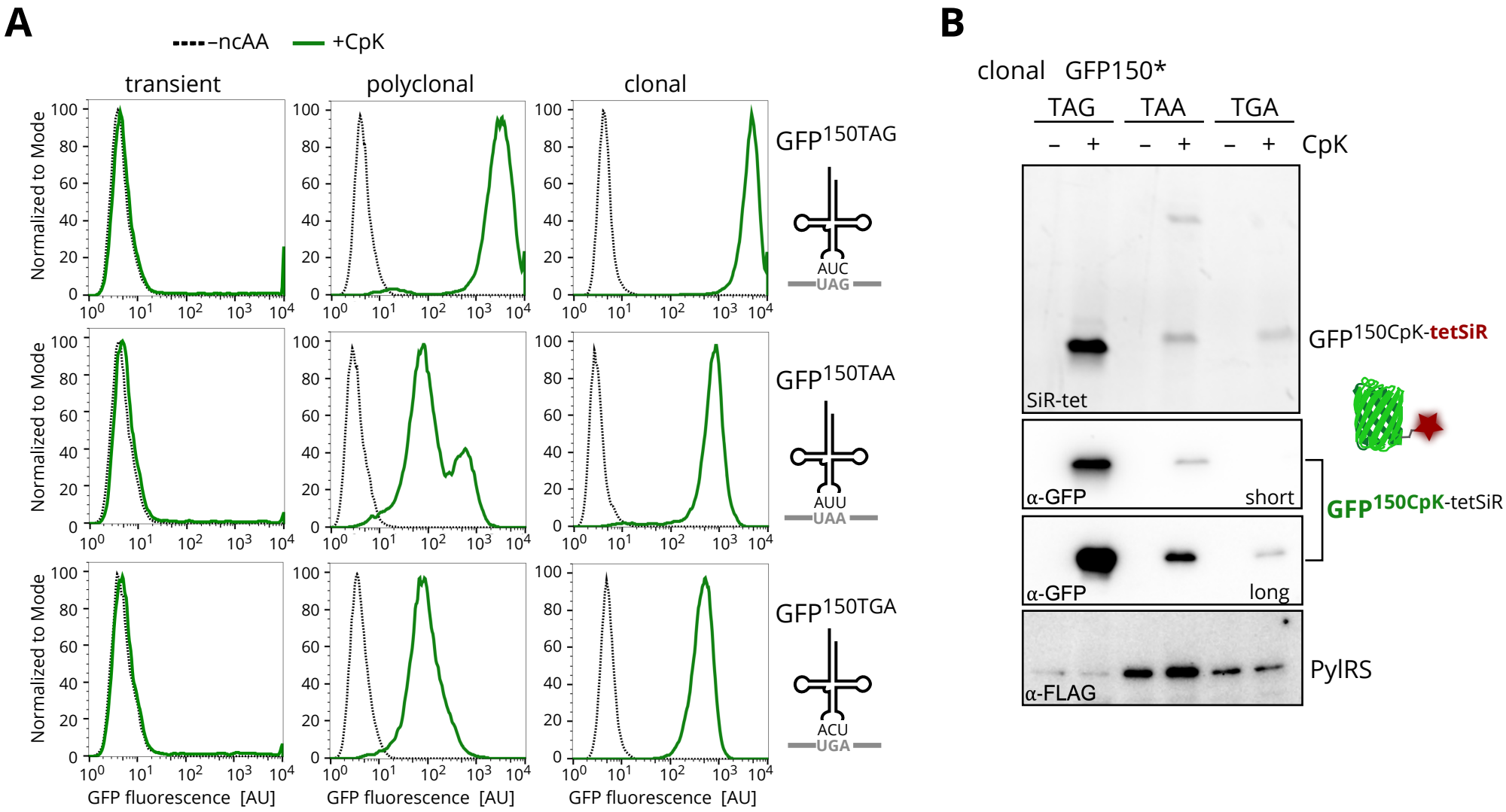

A
–ncAA
+CpK
transient
polyclonal
clonal
Normalized to Mode
GFP fluorescence [AU]
GFP150TAG
AUC
UAG
GFP150TAA
AUU
UAA
GFP150TGA
ACU
UGA
B
clonal GFP150*
TAG
TAA
TGA
– + – + – + CpK
GFP150CpK-tetSiR
SiR-tet
α-GFP
short
long
GFP150CpK-tetSiR
α-FLAG
PylRS

Figure 3 - Generation of efficient ochre and opal suppressor clonal HCT116 populations by PB transgenesis.

A) Flow cytometry comparing amber (UAG, top), ochre (UAA, center) and opal (UGA, bottom) suppression in HCT116 cells after transient transfection (left) and PB transgenesis (center and right). GFP fluorescence was measured without ncAA (dashed black line) or 24 h after addition of 0.2 mM CpK (green line). Left: Cells were transiently transfected with PylT/RS and the PylT/GFP^{150stop} reporter, suppressor PylTs have AUC, AUU, ACU anticodons to match the respective reporter's UAG, UAA and UGA stop codons. Center: using the same plasmids, both PylT/RS and PylT/GFP^{150stop} reporter pAS constructs were integrated into the genome by PB transgenesis. Right: Clonal populations, isolated from the polyclonal pools (middle) by FACS, after 24 h incubation with 0.2 mM CpK. Best clones are shown for amber, ochre and opal suppression, respectively. Other clones recovered are compared in Figure S3.

B) Immunoblot of lysates from clonal stop codon suppressor HCT116-derived cell lines shown in (A). The asterisk (*) is a placeholder for the stop codon indicated. Clonal populations were grown in presence of 0.5 mM CpK for 72 h. GFP^{150CpK} was SPIEDAC labeled with 1 μM tetSiR in lysate followed by SDS-PAGE and immunostaining for GFP expression and FLAG-aaRS expression.

Figure 4

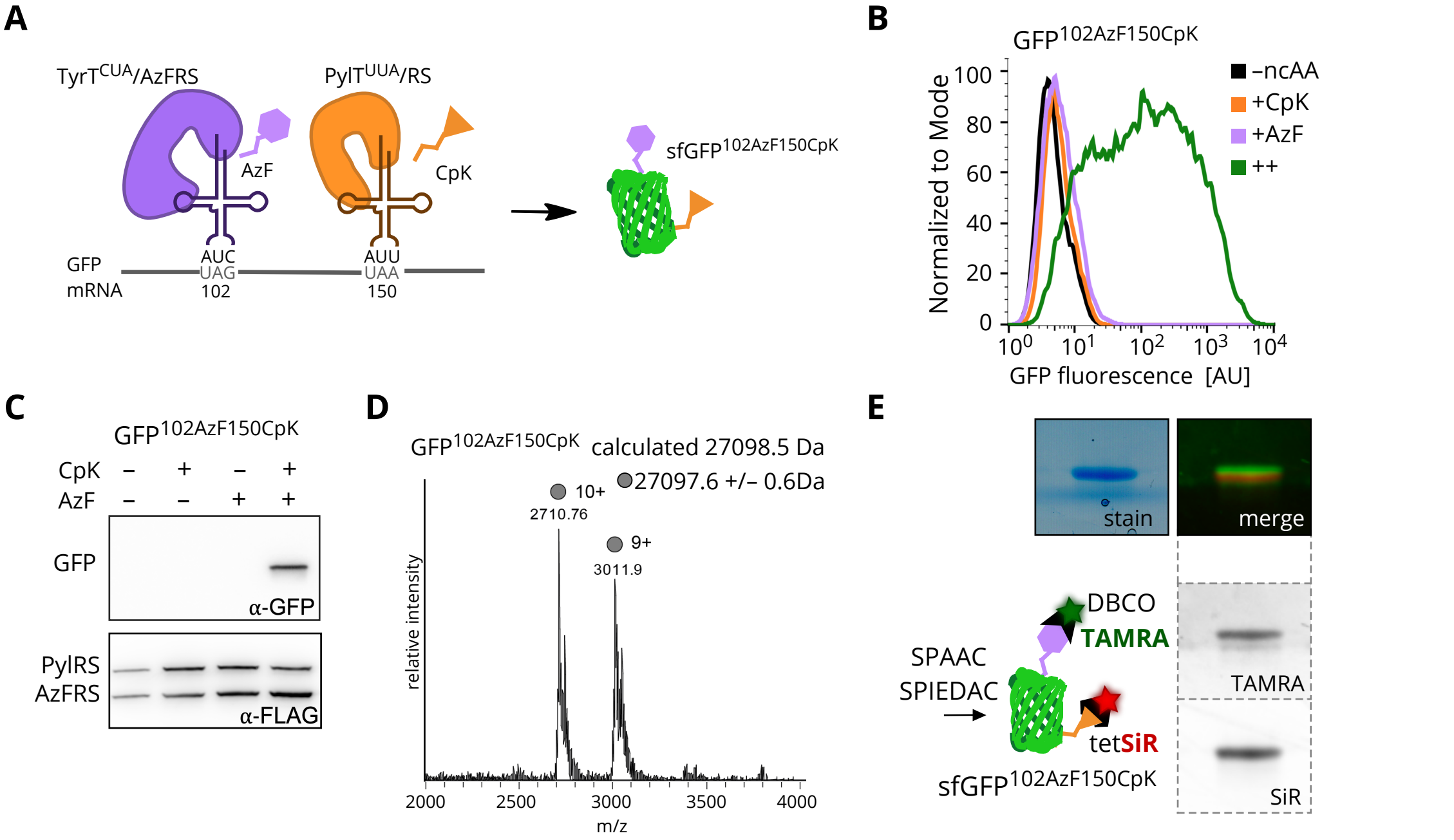

A
TyrT^{CUA}/AzFRS
PylT^{UUA}/RS
AzF
CpK
GFP mRNA
AUC
UAG
102
AUU
UAA
150
sfGFP^{102AzF150CpK}
B
GFP^{102AzF150CpK}
Normalized to Mode
GFP fluorescence [AU]
–ncAA
+CpK
+AzF
++
C
GFP^{102AzF150CpK}
CpK – + – +
AzF – – + +
GFP
α-GFP
PylRS
AzFRS
α-FLAG
D
GFP^{102AzF150CpK} calculated 27098.5 Da
27097.6 +/– 0.6Da
10+
2710.76
9+
3011.9
relative intensity
m/z
E
stain
merge
DBCO
TAMRA
SPAAC
SPIEDAC
tetSiR
sfGFP^{102AzF150CpK}
TAMRA
SiR

Figure 4 - Dual suppression clonal cell line for $GFP^{102AzF150CpK}$ expression in HCT116 cells.

A) Scheme illustrating amber and ochre co-suppression in HCT116 cells with $PylT^{UUA}$/PylRS ochre suppressor pair, $TyrT^{CUA}$/AzFRS amber suppressor pair and a $GFP^{102TAG150TAA}$ reporter integrated by PB transgenesis. Addition of AzF and CpK allows $GFP^{102AzF150CpK}$ production.

B) Flow cytometry of the HCT116 $PylT^{UUA}$/PylRS, $TyrT^{CUA}$/AzFRS and $GFP^{102TAG150TAA}$ integrant clone F6 grown without ncAA (–ncAA, black), with 0.2 mM CpK (+CpK, orange), with 0.5 mM AzF (+AzF, violet) or both ncAAs (++, green) for 48 h.

C) Immunoblot for lysates prepared from the $GFP^{102TAG150TAA}$ integrant clone F6 grown without ncAA, with 0.2 mM CpK, with 0.5 mM AzF or both ncAAs for 48 h. Lysates were separated by SDS-PAGE followed by immunostaining for GFP expression and FLAG-aaRS expression.

D) Intact mass determination for purified $GFP^{102AzF150CpK}$ expressed in the dual suppressor clonal cells. Calculated theoretical mass and determined mass are indicated.

E) $GFP^{102AzF150CpK}$ purified from the PB dual suppressor HCT116 clone by SPAAC and SPIEDAC with 1 µM DBCO-TAMRA and 1 µM tetSiR, respectively. Fluorescence was detected at 520 nm (TAMRA) and 630 nm (SiR) after separation by SDS-PAGE, fluorescence merge and coomassie-stained gel are shown.

Figure 5

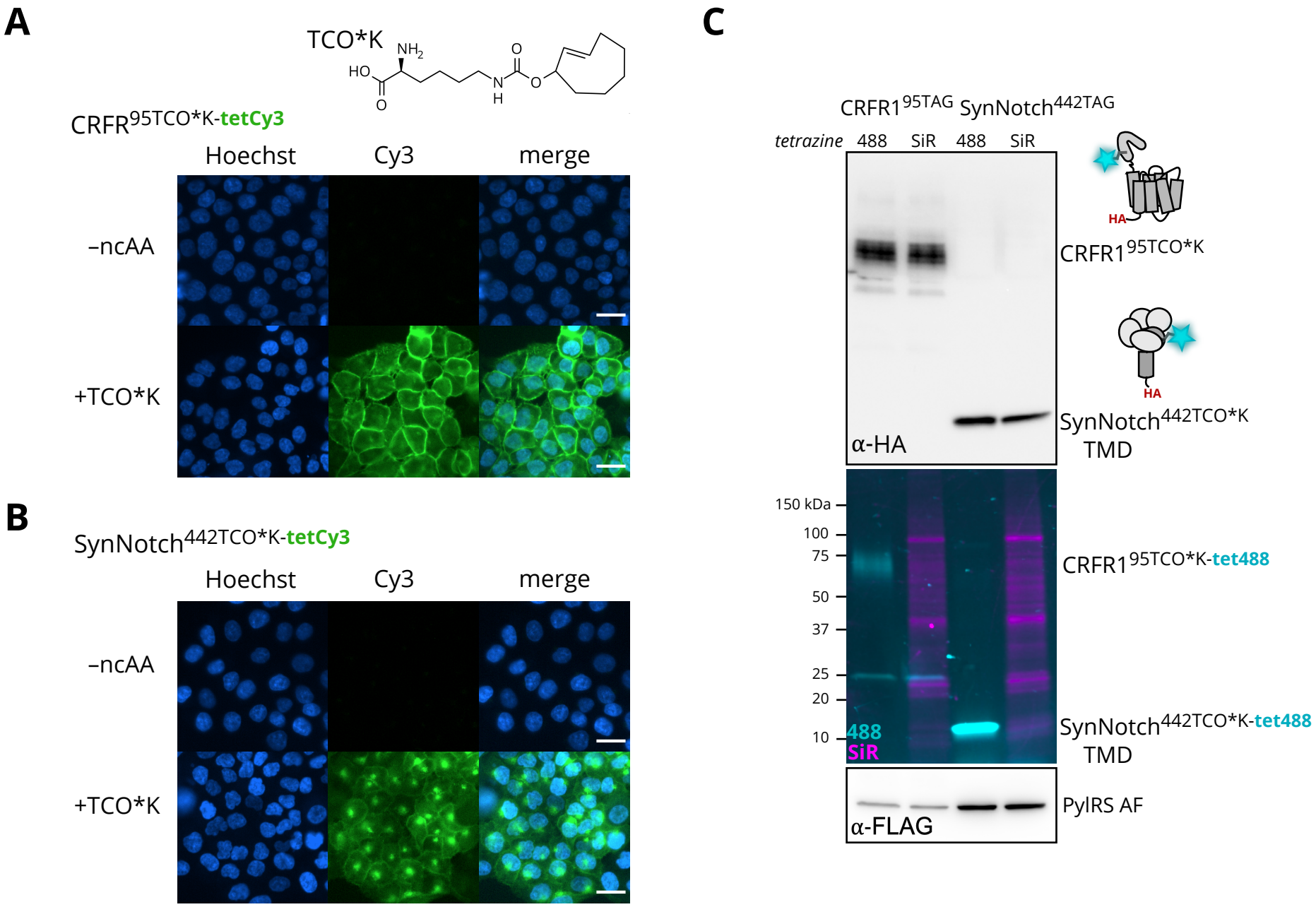

A
TCO*K
NH2
HO
O
N
H
O
O
CRFR95TCO*K-tetCy3
Hoechst
Cy3
merge
–ncAA
+TCO*K
B
SynNotch442TCO*K-tetCy3
Hoechst
Cy3
merge
–ncAA
+TCO*K
C
CRFR195TAG SynNotch442TAG
tetrazine 488 SiR 488 SiR
HA
CRFR195TCO*K
HA
SynNotch442TCO*K
TMD
α-HA
150 kDa
100
75
50
37
25
20
10
488
SiR
CRFR195TCO*K-tet488
SynNotch442TCO*K-tet488
TMD
α-FLAG
PylRS AF

Figure 5 - Amber suppression cell lines for bioorthogonal labeling of cell surface receptors.

Fluorescence microscopy imaging of HCT116 PB integrant cell lines for cell surface receptor expression by amber suppression with (A) PylT/RS AF and PylT/CRFR1^{95TAG} or (B) PylT/RS AF and PylT/SynNotch^{442TAG} integrated by PB transposition. Integrant populations recovered after selection were grown in presence of 0.1 mM TCO*K for 48 h, chased in ncAA free medium for 2 h prior to SPIEDAC labeling with 1 µM tetCy3. Cells were fixed in PFA and nuclei were counterstained with Hoechst 33342 prior to imaging on a Nikon Ti2. White scale bars indicate 25 µm.

C) Immunoblot for the same cell lines as in A) and B) expressing CRFR1^{95TCO*K} or SynNotch^{442TCO*K}. Cells were grown in presence of 0.1 mM TCO*K for 48 h. The cell surface receptors were either labeled by SPIEDAC with 2 µM membrane impermeable tet488 before lysis or with 1 µM tet-SiR in cell lysate. Cleared lysate aliquots were separated by SDS-PAGE and tet488 surface fluorescence and tetSiR lysate labeling visualized at 460 nm (488, cyan) and 630 nm (SiR, magenta), respectively. Immunostaining for the C-terminal HA-tags confirms specific SPIEDAC surface labeling of CRFR1^{95TCO*K} and SynNotch^{N442TCO*K}. Immunostaining for FLAG-PylRS AF is shown. The SynNotch illustration was simplified by omitting the N-terminal LaG17 fusion.

Figure 6

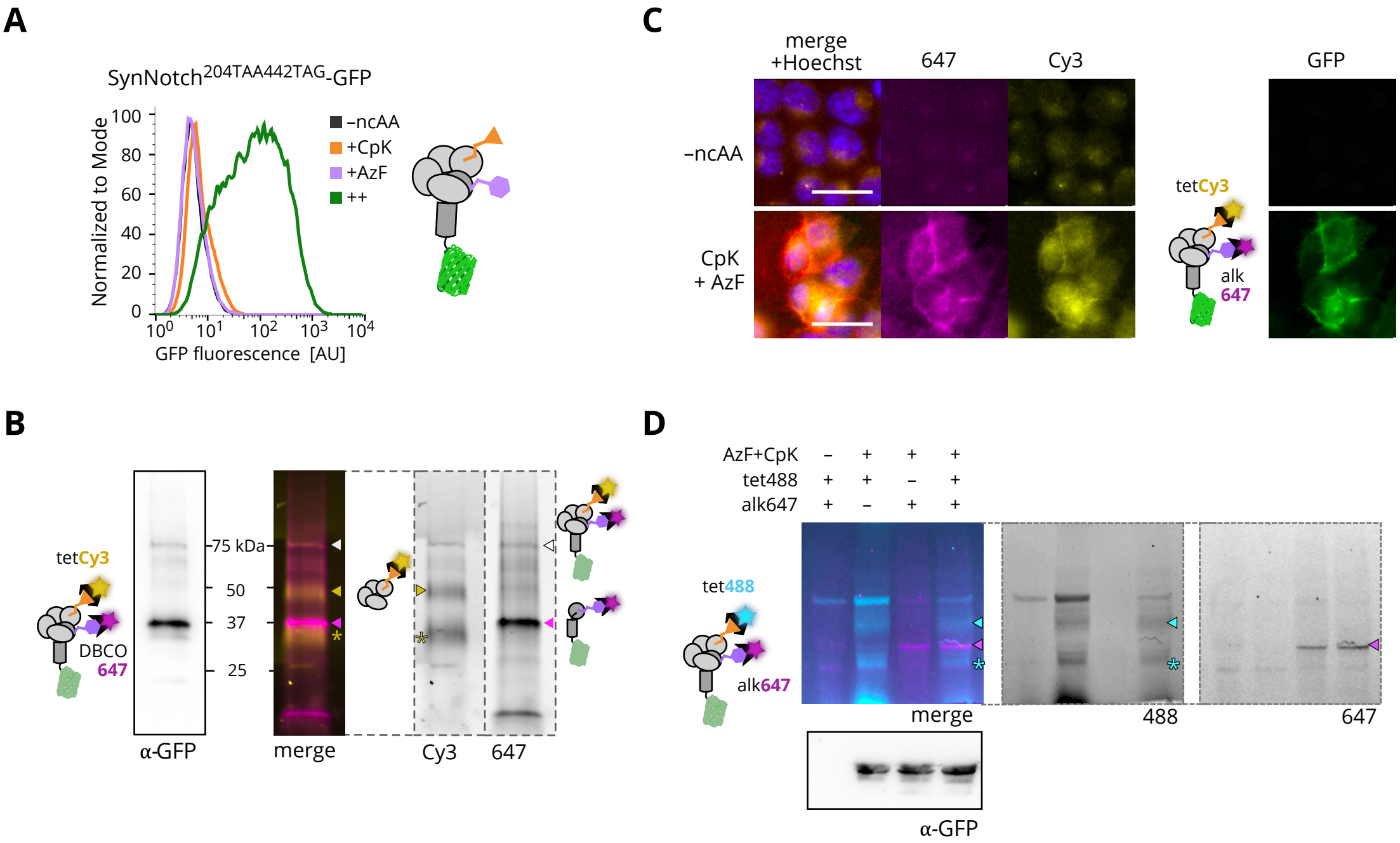

A
SynNotch204TAA442TAG-GFP
Normalized to Mode
100
80
60
40
20
0
10^0 10^1 10^2 10^3 10^4
GFP fluorescence [AU]
–ncAA
+CpK
+AzF
++
C
merge +Hoechst
647
Cy3
GFP
–ncAA
CpK + AzF
tetCy3
alk 647
B
tetCy3
DBCO 647
75 kDa
50
37
25
α-GFP
merge
Cy3
647
D
AzF+CpK – + + +
tet488 + + – +
alk647 + – + +
tet488
alk647
merge
488
647
α-GFP

Figure 6 - Dual suppression HCT116 clonal cell line for SynNotch^{204CpK442AzF} expression and live-cell dual bioorthogonal labeling.

A) The dual suppressor clonal HCT cell line with M15^{UUA}/PylRS ochre suppressor pair, TyrT^{CUA}/AzFRS amber suppressor pair and SynNotch^{204TAA442TAG}-GFP integrated via PB transgenesis was grown 72 h without ncAA (–ncAA, black), 0.2 mM CpK (+CpK, orange), 0.5 mM AzF (+AzF, violet) or both ncAAs (++, green) and analyzed by flow cytometry for GFP fluorescence. The positions of both ncAAs in the domains of the mature receptor are illustrated in the schematic. All SynNotch illustrations were simplified by omitting the N-terminal LaG17 fusion.

B) SynNotch^{204CpK442AzF}-GFP purified from the dual suppression HCT116 integrant clone D8. Bioorthogonal labeling by SPAAC and SPIEDAC with 2 µM DBCO647 and 2 µM tetCy3, respectively. Fluorescence of the bead-captured protein after separation by SDS-PAGE at 520 nm (Cy3), 630 nm (647) and merged signals for in-gel fluorescence, as well as immunoblotting against the C-terminal GFP are shown. The different domains and processing states of SynNotch are indicated - white arrow: 81 kDa pre-processed SynNotch^{204CpK442AzF}; yellow arrow: 43 kDa N-terminal domain with 204CpK labeled by tetCy3; magenta arrow: 38 kDa C-terminal TMD-GFP fusion with 442AzF labeled by DBCO647; yellow asterisk: additional SynNotch N-terminal fragment after proteolytic cleavage of LaG17 (Meineke et al. 2020).

C) Fluorescence microscopy for SynNotch^{204CpK442AzF}-GFP expression in the clonal dual suppressor cell line. Cells were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h, labeled with 2 µM tetCy3 by SPIEDAC and 2 µM alk647 by CuAAC, respectively. Prior to imaging on a Nikon Ti2, nuclei were counterstained with Hoechst 33342. White scale bar indicates 25 µm. See also Figure S6E.

D) SPIEDAC and CuAAC labeling with 2 µM tet488 and 2 µM alk647 on live HCT116 clonal SynNotch^{204CpK442AzF}-GFP cells grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h. Lysates were separated by SDS-PAGE. In-gel fluorescence captured at 460 nm (488) and 630 nm (647) and immunoblotting against the C-terminal GFP are shown. Arrows indicate the different processing states of SynNotch: Cyan arrow indicates the 43 kDa N-terminal domain which contains 204CpK-tet488 after SPIEDAC, magenta arrow indicates 38 kDa C-terminal TMD-GFP fusion which comprising 442AzF-alk647, asterisk indicates an additional SynNotch N-terminal fragment without LaG17 (Meineke et al. 2020).

MATERIAL AND METHODS

DNA constructs

The *M. mazei* PylT/RS expression and super folder GFP^{150TAG} ($sfGFP^{150TAG}$ and sfGFP, referred to as GFP^{150TAG} and GFP throughout) reporter constructs with four tandem h7SK PylT repeats have been described previously (Meineke et al. 2018). Analogous constructs for AnapRS and AzFRS with the respective cognate $LeuT^{CUA}$ or $TyrT^{CUA}$ were generated for this study. The plasmids share a common architecture and are here collectively referred to as "pAS" (Amber Suppression) plasmids: The aaRS, reporter or gene of interest are controlled by EF1α promoter and followed by a internal ribosome entry site (IRES) that allows expression of a downstream selection marker. A cassette with four tandem repeats of the tRNA gene, controlled by the human 7SK PolIII promoter, are placed upstream of the EF1α promoter in anti-sense orientation. A schematic of this architecture is shown in Figure 1A. All DNA constructs were verified by Sanger sequencing and most plasmids are available through Addgene.

Cell Culture and Transfection

HEK293, HEK293T, HCT116, A375, COS-7 and U-2 OS cells were maintained in Dulbecco's modified Eagle's medium (DMEM, GlutaMAX, Thermo) supplemented with 10 % (v/v) FBS (Sigma) at 37 °C and 5% CO_2 atmosphere. For transfection (1.5-2.0) x 10^5 cells/ml were seeded 24 h before transient transfection with 1 µg plasmid DNA/ml using TransIT-LT1 (Mirus) according to the manufacturer's instructions. In transient transfection experiments, ncAAs were added at the time of transfection and cells were harvested after 24, 48 or 72 h, as indicated.

Generation of PB-mediated stable integration cell lines

Cell lines with PiggyBac (PB) integrated transgenes were generated as described (Elsässer et al. 2016; Elsässer 2018): Parental cells were seeded one day prior to transfection and transfected with desired pAS plasmids and PBT in a 4:1 ratio. Earliest 48 h after transfection, cells were split and selected with different concentrations of appropriate antibiotics for 7 days (ranges: 2-5 µg/ml puromycin (VWR), 500-4000 µg/ml blasticidin (Invivogen), 2000-3000 µg/ml geneticin (G418, Sigma)). Stable integrant cell populations were recovered from the highest submissible selection condition in DMEM + 10% FBS (v/v). Polyclonal populations were grown in presence of ncAA for 24, 48 or 72 h, as indicated, and characterized by flow cytometry for reporter fluorescence and immunoblotting for reporter or protein of interest and PylRS. A list of

cell lines generated in this study can be found on Mendeley data (DOI: 10.17632/tgc7mbv5xp.1).

Noncanonical amino acids

Working stocks of ncAAs were prepared in 100 mM NaOH with 15% DMSO (v/v) to 100 mM. Anap working stock was prepared to 20 mM in the same buffer. Chemical structures of N-ε-[(2-methyl-2-cyclopropene-1-yl)-methoxy] carbonyl-L-lysine (CpK), p-azido-phenylalanine (AzF), 3-(6-acetylnaphthalen-2-ylamino)-2-amino-propanoic acid (Anap) and axial trans-cyclooct-2-ene-L-lysine (TCO*K) are shown in Figures 1, 2 and 5, respectively. Refer to the Key Resource Table for supplier and CAS numbers.

SPIEDAC lysate labeling, SDS-PAGE and immunoblotting

Aliquots of cell lysate were separated on 4-20 % Tris-glycine gels (BioRad) and transferred to nitrocellulose membranes (BioRad). Cells were lysed in RIPA (for intracellular proteins) or PBS with 0.1% triton X-100 (v/v) (cell membrane proteins) supplemented with 1x cOmplete proteinase inhibitor (Roche). The insoluble fraction was removed by centrifugation. To compare amber suppression of PB transgenic integrants generated from different human cells, the soluble fraction of whole cell lysates were normalized for total protein content (BCA assay, Pierce). For SPIEDAC labeling in lysate, 1 μM tetrazine-Silicon rhodamine (tetSiR) (Spirochrome) was added to the lysis buffer. Lysates cleared by centrifugation were mixed with 6x Laemmli buffer and denatured for 10 min at 37°C, or at 95°C to completly denature GFP. After separation, gels were exposed at 460 nm, 520 nm and 630 nm in a GE AI600 imager for in-gel fluorescence. Expression of GFP reporter, HA- or GFP-fusion proteins and FLAG-aaRS was confirmed by immunoblotting with antibodies against GFP (Santa Cruz, sc-9996), HA-HRP (Roche, 12013819001), FLAG-HRP (Sigma, A8592), β-actin (cell signaling, 4970), GAPDH (Millipore, AB2302) and corresponding secondary HRP-conjugated antibodies when needed (BioRad). Annotated, preprocessed immunoblotting data can be found on Mendeley data (DOI: 10.17632/tgc7mbv5xp.1).

Quantification of GFP expression

HEK293T-derived cells, selected for pAS_4x*Mma*PylT/RS and pAS_4x*Mma*PylT/GFP^{150TAG} PB transposition (low antibiotic concentration: 2 μg/ml puromycin, 500 μg/ml blasticidin), were

transiently transfected with an additional pAS plasmid in quadruplicate. Transfected cells were grown in absence or presence (triplicate) of 0.2 mM CpK for 24 h. Cells were lysed in RIPA buffer with 1x cOmplete protease inhibitor (Roche), the insoluble fraction was removed by centrifugation. GFP bottom fluorescence of an aliquot was measured in Tecan Infinite M200 pro plate reader (excitation 485 nm, emission 518 nm). Fluorescence measurements were normalized to total protein content of each sample as determined by Pierce BCA assay kit (Fisher Scientific) on the same sample. Average GFP fluorescence was calculated from the technical triplicates, the GFP fluorescence measured in absence of CpK was subtracted as background fluorescence and values were normalized to average GFP fluorescence measured for the untransfected amber suppression cell line.

Flow cytometry and fluorescence-assisted cell sorting (FACS)

Cells were trypsinized after the indicated time of culturing in presence of ncAAs and resuspended in PBS + 10% FBS (v/v). Flow cytometry was carried out on a Beckmann NAVIOS flow cytometer and analyzed in FlowJo™ software (BD Life Science, version 10.6.2).

Polyclonal populations selected with antibiotics after PB transgenesis were grown in the presence of the cognate ncAAs for 24 h (GFP^{150TAG}, GFP^{150TAA}, GFP^{150TGA}), 48 h ($GFP^{102TAG150TAA}$) or 72 h ($SynNotch^{204TAA442TAG}$-GFP). Single cells were isolated by gating for top 0.4-5% GFP fluorescence signal on a SONY SH800 cell sorter. Clonal populations were expanded in DMEM + 10% FBS and analyzed by flow cytometry and immunoblotting for GFP and aaRS expression.

Intact mass spectrometry

The dual suppression $GFP^{102TAG150TAA}$ clonal integrant cell line was cultured in the presence of 0.2 mM and 0.5 mM AzF for 72 h. Cells were lysed in RIPA buffer supplemented with 1x cOmplete protease inhibitor (Roche). The insoluble fraction was removed by centrifugation. Expressed GFP was captured on GFP-Trap_MA magnetic beads (ChromoTEK), washed with RIPA buffer, PBS + 500 mM NaCl and PBS prior to elution in 1 % (v/v) acetic acid.

Purified $GFP^{102AzF150CpK}$ samples were desalted and rebuffered into 100 mM ammonium acetate pH 7.5 using ZebaSpin columns with a 7 kDa cut-off (Thermo). Samples were directly infused into an Orbitrap Fusion Tribrid mass spectrometer equipped with an offline nanospray source using borosilicate capillaries (Thermo). The capillary voltage was 1.5 kV and the pressure in the ion-routing multipole was maintained at 0.11 torr. Spectra were acquired in the Orbitrap mass

analyzer operated in high mass mode at a resolution of 60.000 between 800-3000 *m/z*. Data were analyzed using Excalibur (Thermo).

Bioorthogonal labeling on beads

$GFP^{102AzF150CpK}$ and $SynNotch^{204CpK442AzF}$-GFP were captured on GFP-Trap_MA magnetic beads (ChromoTEK) from cell lysate of the respective PB stable clonal cell lines cultured in presence of 0.5 mM AzF and 0.2 mM CpK for 72 h. Cells were lysed and the insoluble fraction was removed by centrifugation. Beads were washed with RIPA buffer, PBS + 500 mM NaCl and PBS. SPIEDAC and SPAAC were carried out on bead-bound proteins simultaneously by adding 1 µM tetSiR (spirochrome) or 1 µM tetCy3 (Click Chemistry Tools) and 1 µM DBCO-TAMRA (Jena Bioscience) or 1 µM DBCO647 (Jena Bioscience) for 10 min on ice. Excess dye was washed off and bead-bound proteins were eluted with 1 % (v/v) acetic acid for SDS-PAGE separation. Equal amounts of GFP purified from different samples were separated on 4-20% Tris-glycine gels (BioRad) and exposed for in-gel fluorescence at 520 nm and 630 nm in a GE AI600 imager. The gel was stained with InstantBlue (Expedeon) to visualize $GFP^{102AzF150CpK}$ bands or transferred to a nitrocellulose membrane for immunostaining against GFP to detect $SynNotch^{204CpK442AzF}$-GFP.

Live-cell surface SPIEDAC and CuAAC labeling for in-gel fluorescence and fluorescence microscopy

Surface SPIEDAC labeling was performed with 1-2 µM tet488 (Jena Bioscience) or 1-2µM tetCy3 (Click Chemistry Tools) in DMEM + 10% FBS (v/v) for 30 min at 37 °C. Where indicated SPAAC was carried out with 2 µM DBCO-AF647 (DBCO647, Click Chemistry Tools) in DMEM +10% FBS (v/v) for 30 min at 37 °C. For subsequent CuAAC in the dual suppression cell line, cells were washed with PBS and labeling was performed with 50 µM $CuSO_4$, 250 µM THPTA, 2.5 mM sodium ascorbate and 2 µM alk647 (Jena Bioscience) for 10 min at 22 °C. Cells were washed and collected in PBS and lysed in PBS with 0.2% triton X-100 (v/v) and 1x complete protease inhibitor (Roche) on ice. Lysates cleared by centrifugation were mixed with 6x Laemmli buffer and denatured for 10 min at 37°C, or at 95°C to completly denature GFP. Proteins were separated by SDS-PAGE and in-gel fluorescence measured as described above. Alternatively, for cells grown in poly-L-lysine coated 18-well glass imaging slides (Ibidi), cells were washed with PBS, fixed in paraformaldehyde (PFA) 4% (v/v) and counterstained with 2 µM hoechst 33342 prior to imaging in PBS on a Nikon Eclipse Ti2 inverted widefield microscope, using a 20x

(0.75 NA) objective and filter sets for DAPI, GFP, Cy3 and Cy5 fluorescence. Images were processed in Fiji (Schindelin et al. 2012).

Table 1 - Reagents and resources

Antibodies		
Mouse monoclonal anti-GFP	Santa Cruz	Cat#sc-9996; RRID:AB_627695
Rat monoclonal anti-HA HRP-coupled	Roche	Cat#12013819001; RRID:AB_390917
Mouse monoclonal anti-FLAG HRP-coupled	Sigma	Cat#A8592; RRID:AB_439702
Rabbit monoclonal anti-β-actin	Cell Signaling	Cat#4970; RRID:AB_2223172
Chicken polyclonal GAPDH	Millipore	Cat#AB2302; RRID:AB_10615768
Goat anti-chicken IgY (H+L) HRP	Invitrogen	Cat#A16054; RRID:AB_2534727
Goat anti-mouse HRP-coupled	BioRad	Cat#1721011; RRID:AB_11125936
Goat anti-rabbit HRP-coupled	BioRad	Cat#1721019; RRID:AB_11125143
Chemicals, Peptides, and Recombinant Proteins		
CpK: N-ε-[(2-methyl-2-cyclopropene-1-yl)-methoxy] carbonyl-L-lysine	SiChem	Cat#SC-8017; CAS:1610703-09-7
ProK: N-propargyl-L-lysine	Iris Biotech	Cat#HAA2090; CAS:1428330-91-9
TCO*K: axial trans-cyclooct-2-ene-L-lysine	SiChem	Cat#SC-8008; CAS:1801936-26-4
Anap: (2S)-3-[(6-Acetylnaphthalen-2-yl)amino]-2-aminopropanoic Acid	Cayman	Cat#15436; CAS 1313516-26-5
AzF: 4-Azido-L-phenylalanine	Santa Cruz	Cat#sc-289923; CAS: 33173-53-4
tetrazine-Siliconrhodamine (tet-SiR)	Spirochrom	Cat#SC008
tetrazine-AF488 (tet488)	Click Chemistry Tools	Cat#1361
Dibenzylcyclooctyne-PEG4-5/6-Tetramethyl rhodamine (DBCO-TAMRA)	Jena Bioscience	Cat#CLK-A131
Cy3-tetrazine (tetCy3)	Click Chemistry Tools	Cat#1204
alkyne-AF647 (alk647)	Jena Bioscience	Cat#CLK-1301A
Dibenzylcyclooctyne-AF647 (DBCOAF647)	Jena Bioscience	Cat#CLK-1302
Critical Commercial Assays		
Pierce BCA assay kit	Fisher Scientific	Cat#23227
GFP-Trap_MA magnetic beads	ChromoTek	Cat#gtma
Deposited Data		
Mendeley data	DOI: 10.17632/tgc7mbv5xp.1	
Experimental Models: Cell Lines		

Human: HEK293T cells		N/A
Human: HEK293 cells		N/A
Human: HCT116 cells		N/A
Human: A375		N/A
Human: U-2 OS		N/A
Simian: COS7		N/A
Recombinant DNA		
Piggy Bac Transposase expression plasmid	Sytem Bioscience	Cat#PB510B-1
pAS_*Mma*PylT/PylRS	(Meineke et al. 2018)	Addgene#140009
pAS_*Mma*PylT/GFP150TAG	(Meineke et al. 2018)	Addgene#140015
pAS_*Mma*PylT/GFP150TAG	(Meineke et al. 2020)	Addgene#154766
pAS_sfGFP102TAG133TGA150TAA	this study	N/A
pAS_*Mma*PylRS	(Meineke et al. 2018)	Addgene#154762
pAS (MmaPylT)	(Meineke et al. 2018)	Addgene#14008
pAS_*Mma*PylT/AbKRS	this study	Addgene#140020
pAS_EcoLeuT/AnapRS	this study	Addgene#174894
pAS_EcoLeuT/GFP150TAG	this study	Addgene#174892
pAS_*Mma*PylT(UUA)/PylRS	this study	Addgene#174896
pAS_*Bst*TyrT(CUA)/*Eco*TyrRS	this study	Addgene#140018
pAS_*Bst*TyrT(CUA)/sfGFP150TAG	this study	Addgene#174891
pAS_*Mma*PylT/PylRS AF	(Meineke et al. 2018)	Addgene#140023
pAS_*Bst*TyrT(CUA)/sfGFP102TAG150TAA	this study	Addgene#174895
pAS_*EcoLeu*T(CUA)/AnapRS	this study	Addgene#140019
pAS_*Mma*PylT(UUA)/GFP150TAA	this study	Addgene#174897
pAS_*Mma*PylT(UCA)/GFP150TGA	this study	Addgene#174898
pAS_*Mma*PylT(UCA)/PylRS	this study	Addgene#174899
pAS_M15T(UUA)/*MmaPyl*RS	(Meineke et al. 2020)	Addgene#154774
pAS_*Mma*PylT/CRFR1 95TAG	(Meineke et al. 2021)	Addgene#174900
pAS_*Mma*PylT/LaG17-SynNotch442TAG	this study	N/A
pAS_M15PylT(UUA)/LaG17-SynNotch39TAA277TAG	this study	N/A
Software and Algorithms		
Fiji Schindelin et al., 2012	(Schindelin et al. 2012)	https://fiji.sc; RRID:SCR_002285
FlowJoTM	BD Life Sciences	v10.6.2 Software

REFERENCES

Ai H, Shen W, Sagi A, Chen PR, Schultz PG. 2011. Probing protein-protein interactions with a genetically encoded photo-crosslinking amino acid. *Chembiochem* **12**: 1854–1857.

Ambrogelly A, Gundllapalli S, Herring S, Polycarpo C, Frauer C, Söll D. 2007. Pyrrolysine is not hardwired for cotranslational insertion at UAG codons. *Proc Natl Acad Sci USA* **104**: 3141–3146.

Axup JY, Bajjuri KM, Ritland M, Hutchins BM, Kim CH, Kazane SA, Halder R, Forsyth JS, Santidrian AF, Stafin K, et al. 2012. Synthesis of site-specific antibody-drug conjugates using unnatural amino acids. *Proc Natl Acad Sci USA* **109**: 16101–16106.

Bartoschek MD, Ugur E, Nguyen T-A, Rodschinka G, Wierer M, Lang K, Bultmann S. 2021. Identification of permissive amber suppression sites for efficient non-canonical amino acid incorporation in mammalian cells. *Nucleic Acids Res* **49**: e62.

Beránek V, Willis JCW, Chin JW. 2019. An Evolved Methanomethylophilus alvus Pyrrolysyl-tRNA Synthetase/tRNA Pair Is Highly Active and Orthogonal in Mammalian Cells. *Biochemistry* **58**: 387–390.

Bryson DI, Fan C, Guo L-T, Miller C, Söll D, Liu DR. 2017. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol* **13**: 1253–1260.

Chatterjee A, Guo J, Lee HS, Schultz PG. 2013. A genetically encoded fluorescent probe in mammalian cells. *J Am Chem Soc* **135**: 12540–12543.

Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG. 2003. An expanded eukaryotic genetic code. *Science* **301**: 964–967.

Chin JW. 2014. Expanding and reprogramming the genetic code of cells and animals. *Annu Rev Biochem* **83**: 379–408.

Coin I, Katritch V, Sun T, Xiang Z, Siu FY, Beyermann M, Stevens RC, Wang L. 2013. Genetically encoded chemical probes in cells reveal the binding path of urocortin-I to CRF class B GPCR. *Cell* **155**: 1258–1269.

Ding S, Wu X, Li G, Han M, Zhuang Y, Xu T. 2005. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**: 473–483.

Elliott TS, Townsley FM, Bianco A, Ernst RJ, Sachdeva A, Elsässer SJ, Davis L, Lang K, Pisa R, Greiss S, et al. 2014. Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal. *Nat Biotechnol* **32**: 465–472.

Elsässer SJ, Ernst RJ, Walker OS, Chin JW. 2016. Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat Methods* **13**: 158–164.

Elsässer SJ. 2018. Generation of stable amber suppression cell lines. *Methods Mol Biol* **1728**: 237–245.

Fliedl L, Kast F, Grillari J, Wieser M, Grillari-Voglauer R. 2015. Optimization of a quantitative

PCR based method for plasmid copy number determination in human cell lines. *N Biotechnol* **32**: 716–719.

Floquet C, Hatin I, Rousset J-P, Bidou L. 2012. Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. *PLoS Genet* **8**: e1002608.

Gautier A, Nguyen DP, Lusic H, An W, Deiters A, Chin JW. 2010. Genetically encoded photocontrol of protein localization in mammalian cells. *J Am Chem Soc* **132**: 4086–4088.

Jewel D, Kelemen RE, Huang RL, Zhu Z, Sundaresh B, Cao X, Malley K, Huang Z, Pasha M, Anthony J, et al. 2023. Virus-assisted directed evolution of enhanced suppressor tRNAs in mammalian cells. *Nat Methods* **20**: 95–103.

Lafranchi L, Schlesinger D, Kimler KJ, Elsässer SJ. 2020. Universal Single-Residue Terminal Labels for Fluorescent Live Cell Imaging of Microproteins. *J Am Chem Soc* **142**: 20080–20087.

Lee HS, Guo J, Lemke EA, Dimla RD, Schultz PG. 2009. Genetic incorporation of a small, environmentally sensitive, fluorescent probe into proteins in Saccharomyces cerevisiae. *J Am Chem Soc* **131**: 12921–12923.

Liu W, Brock A, Chen S, Chen S, Schultz PG. 2007. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* **4**: 239–244.

Meineke B, Heimgärtner J, Craig AJ, Landreh M, Moodie LWK, Elsässer SJ. 2021. A genetically encoded picolyl azide for improved live cell copper click labeling. *Front Chem* **9**: 768535.

Meineke B, Heimgärtner J, Eirich J, Landreh M, Elsässer SJ. 2020. Site-Specific Incorporation of Two ncAAs for Two-Color Bioorthogonal Labeling and Crosslinking of Proteins on Live Mammalian Cells. *Cell Rep* **31**: 107811.

Meineke B, Heimgärtner J, Lafranchi L, Elsässer SJ. 2018. Methanomethylophilus alvus Mx1201 Provides Basis for Mutual Orthogonal Pyrrolysyl tRNA/Aminoacyl-tRNA Synthetase Pairs in Mammalian Cells. *ACS Chem Biol* **13**: 3087–3096.

Muerdter F, Boryń ŁM, Woodfin AR, Neumayr C, Rath M, Zabidi MA, Pagani M, Haberle V, Kazmar T, Catarino RR, et al. 2018. Resolving systematic errors in widely used enhancer activity assays in human cells. *Nat Methods* **15**: 141–149.

Nikić I, Plass T, Schraidt O, Szymański J, Briggs JAG, Schultz C, Lemke EA. 2014. Minimal tags for rapid dual-color live-cell labeling and super-resolution microscopy. *Angew Chem Int Ed* **53**: 2245–2249.

Niu W, Schultz PG, Guo J. 2013. An expanded genetic code in mammalian cells with a functional quadruplet codon. *ACS Chem Biol* **8**: 1640–1645.

Oller-Salvia B, Kym G, Chin JW. 2018. Rapid and Efficient Generation of Stable Antibody-Drug

Conjugates via an Encoded Cyclopropene and an Inverse-Electron-Demand Diels-Alder Reaction. *Angew Chem Int Ed* **57**: 2831–2834.

Osgood AO, Zheng Y, Roy SJS, Loynd C, Jewel D, Chatterjee A. 2022. An efficient opal-suppressor tryptophanyl pair creates new routes for simultaneously incorporating up to three distinct noncanonical amino acids into proteins in mammalian cells. *BioRxiv*. doi: 10.1101/2022.08.02.502539.

Palomar-Siles M, Heldin A, Zhang M, Strandgren C, Yurevych V, van Dinter JT, Engels SAG, Hofman DA, Öhlin S, Meineke B, et al. 2022. Translational readthrough of nonsense mutant TP53 by mRNA incorporation of 5-Fluorouridine. *Cell Death Dis* **13**: 997.

Rother M, Krzycki JA. 2010. Selenocysteine, pyrrolysine, and the unique energy metabolism of methanogenic archaea. *Archaea* **2010**.

Sakamoto K, Hayashi A, Sakamoto A, Kiga D, Nakayama H, Soma A, Kobayashi T, Kitabatake M, Takio K, Saito K, et al. 2002. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res* **30**: 4692–4699.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.

Schmied WH, Elsässer SJ, Uttamapinant C, Chin JW. 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.

Serfling R, Lorenz C, Etzel M, Schicht G, Böttke T, Mörl M, Coin I. 2018. Designer tRNAs for efficient incorporation of non-canonical amino acids by the pyrrolysine system in mammalian cells. *Nucleic Acids Res* **46**: 1–10.

Shao S, Koh M, Schultz PG. 2020. Expanding the genetic code of the human hematopoietic system. *Proc Natl Acad Sci USA* **117**: 8845–8849.

Srinivasan G, James CM, Krzycki JA. 2002. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**: 1459–1462.

Suzuki T, Miller C, Guo L-T, Ho JML, Bryson DI, Wang Y-S, Liu DR, Söll D. 2017. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat Chem Biol* **13**: 1261–1266.

van Geel R, Pruijn GJM, van Delft FL, Boelens WC. 2012. Preventing thiol-yne addition improves the specificity of strain-promoted azide-alkyne cycloaddition. *Bioconjug Chem* **23**: 392–398.

van Husen LS, Katasori A-M, Meineke B, Tjernberg LO, Schedin-Weiss S, Elsässer SJ. 2021. Engineered human induced pluripotent cells enable genetic code expansion in brain organoids. *Chembiochem*. doi: 10.1002/cbic.202100399.

Wan W, Tharp JM, Liu WR. 2014. Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool. *Biochim Biophys Acta* **1844**: 1059–1070.

Xiao H, Chatterjee A, Choi S, Bajjuri KM, Sinha SC, Schultz PG. 2013. Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew Chem Int Ed* **52**: 14080–14083.

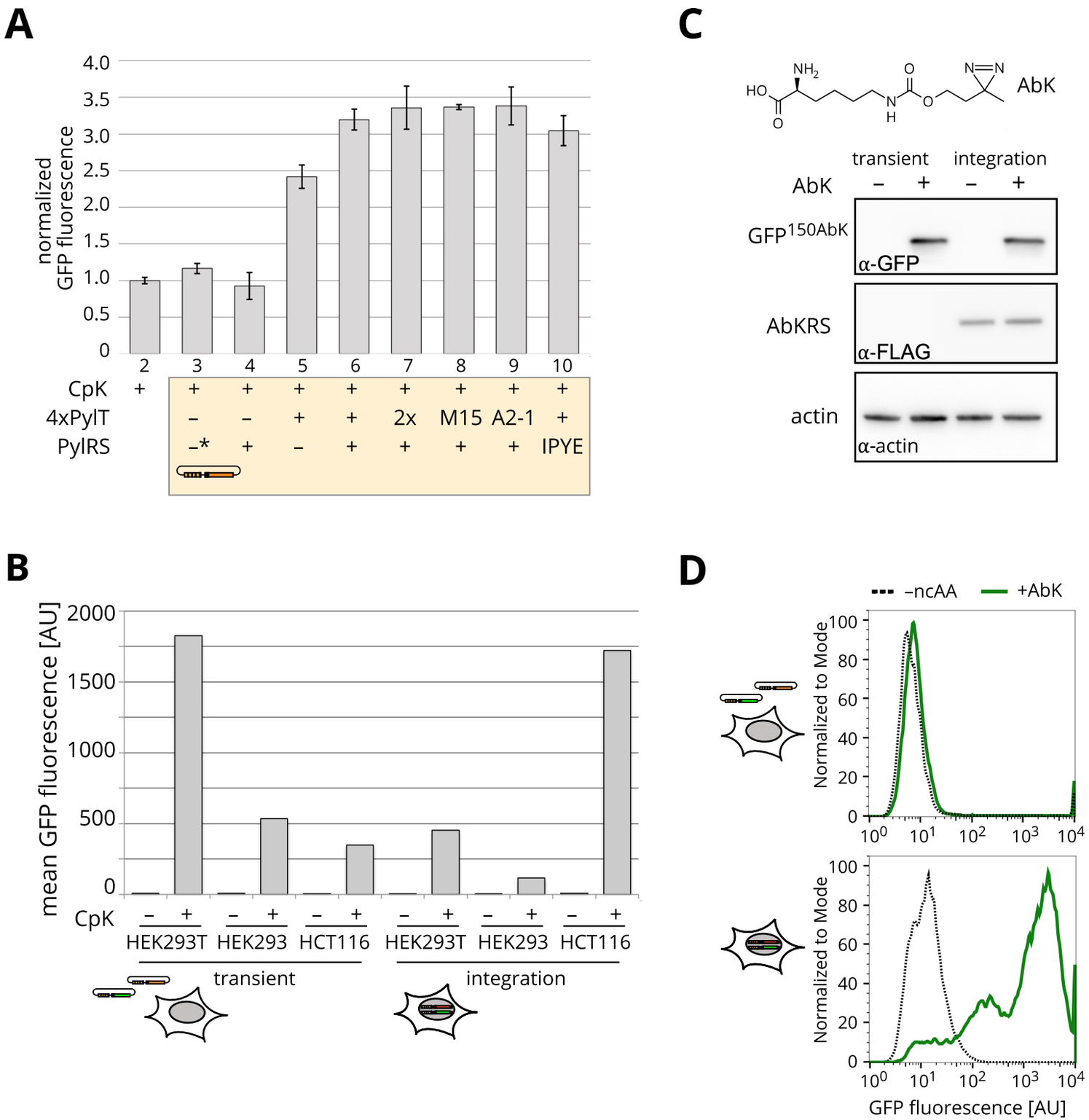
Xi Z, Davis L, Baxter K, Tynan A, Goutou A, Greiss S. 2022. Using a quadruplet codon to expand the genetic code of an animal. *Nucleic Acids Res* **50**: 4801–4812.

Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, Yokoyama S. 2008. Multistep engineering of pyrrolysyl-tRNA synthetase to genetically encode N(epsilon)-(o-azidobenzyloxycarbonyl) lysine for site-specific protein modification. *Chem Biol* **15**: 1187–1197.

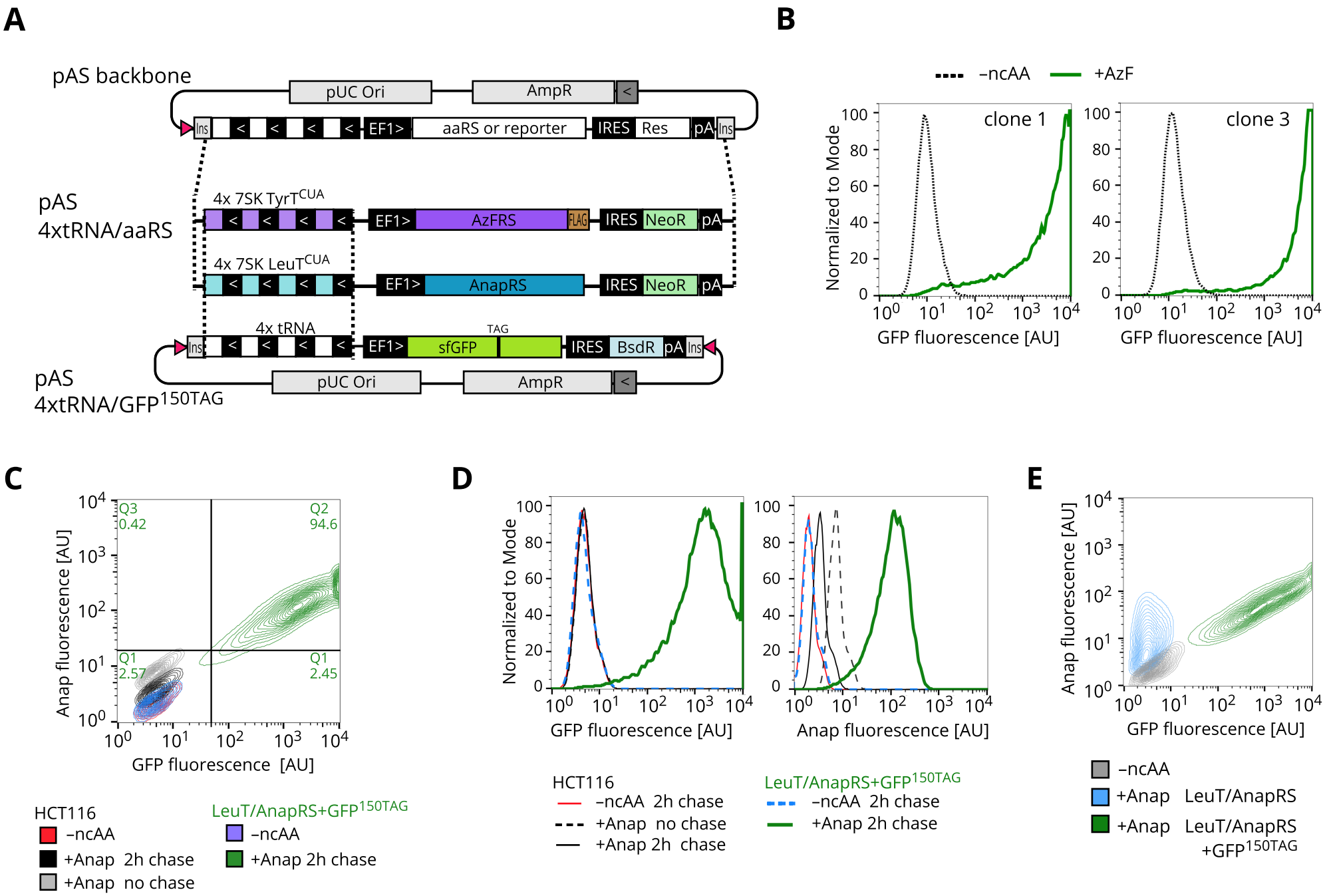
Zheng Y, Addy PS, Mukherjee R, Chatterjee A. 2017. Defining the current scope and limitations of dual noncanonical amino acid mutagenesis in mammalian cells. *Chem Sci* **8**: 7211–7217.

Zhou W, Wesalo JS, Liu J, Deiters A. 2020. Genetic code expansion in mammalian cells: A plasmid system comparison. *Bioorg Med Chem* **28**: 115772.

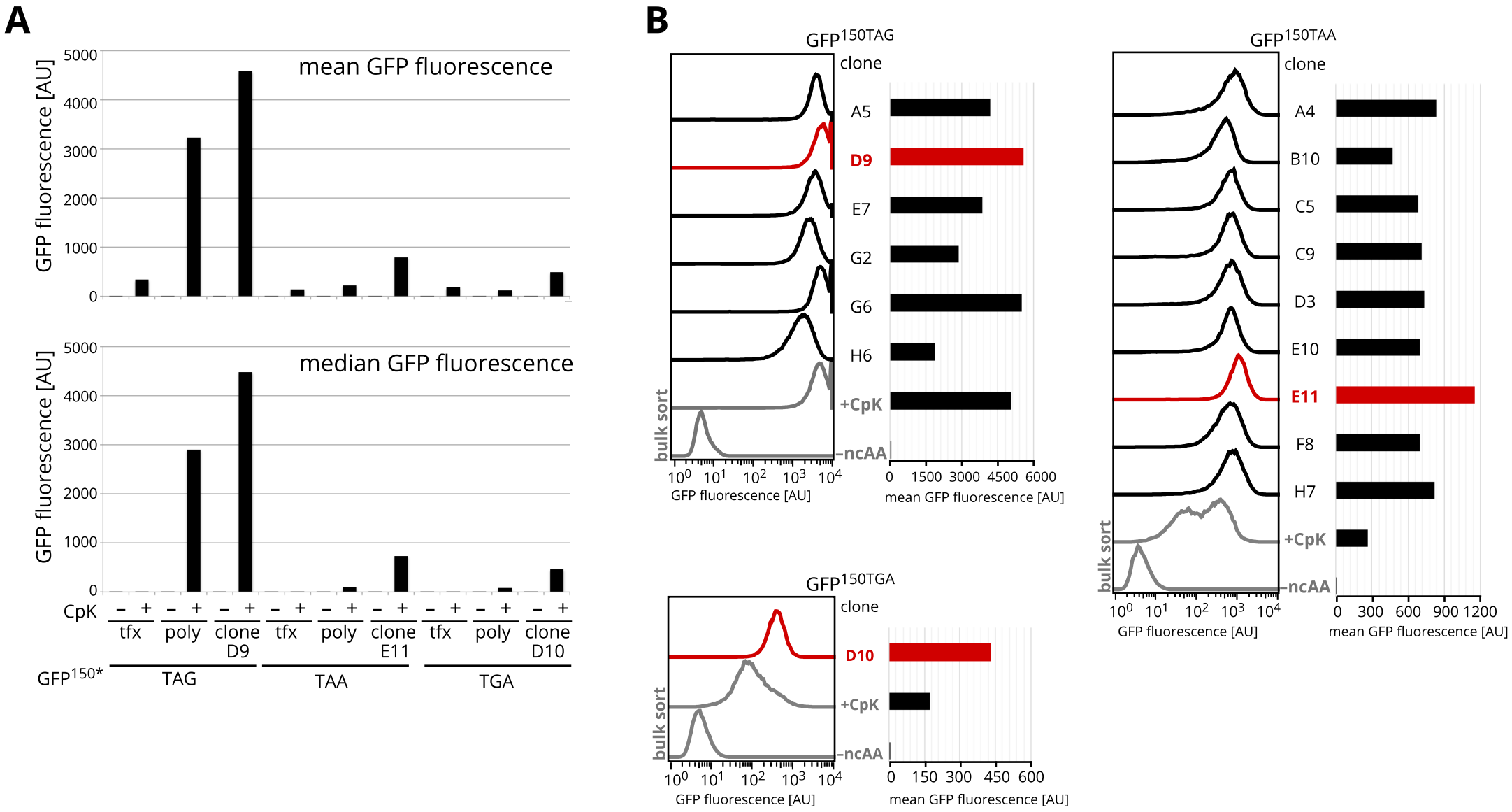
to Figure 1


A
normalized GFP fluorescence
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
2 3 4 5 6 7 8 9 10
CpK + + + + + + + + +
4xPylT – – + + 2x M15 A2-1 +
PylRS –* + – + + + + IPYE
B
mean GFP fluorescence [AU]
2000
1500
1000
500
0
CpK – + – + – + – + – + – +
HEK293T HEK293 HCT116 HEK293T HEK293 HCT116
transient
integration
C
AbK
transient integration
AbK – + – +
GFP150AbK
α-GFP
AbKRS
α-FLAG
actin
α-actin
D
–ncAA
+AbK
Normalized to Mode
GFP fluorescence [AU]

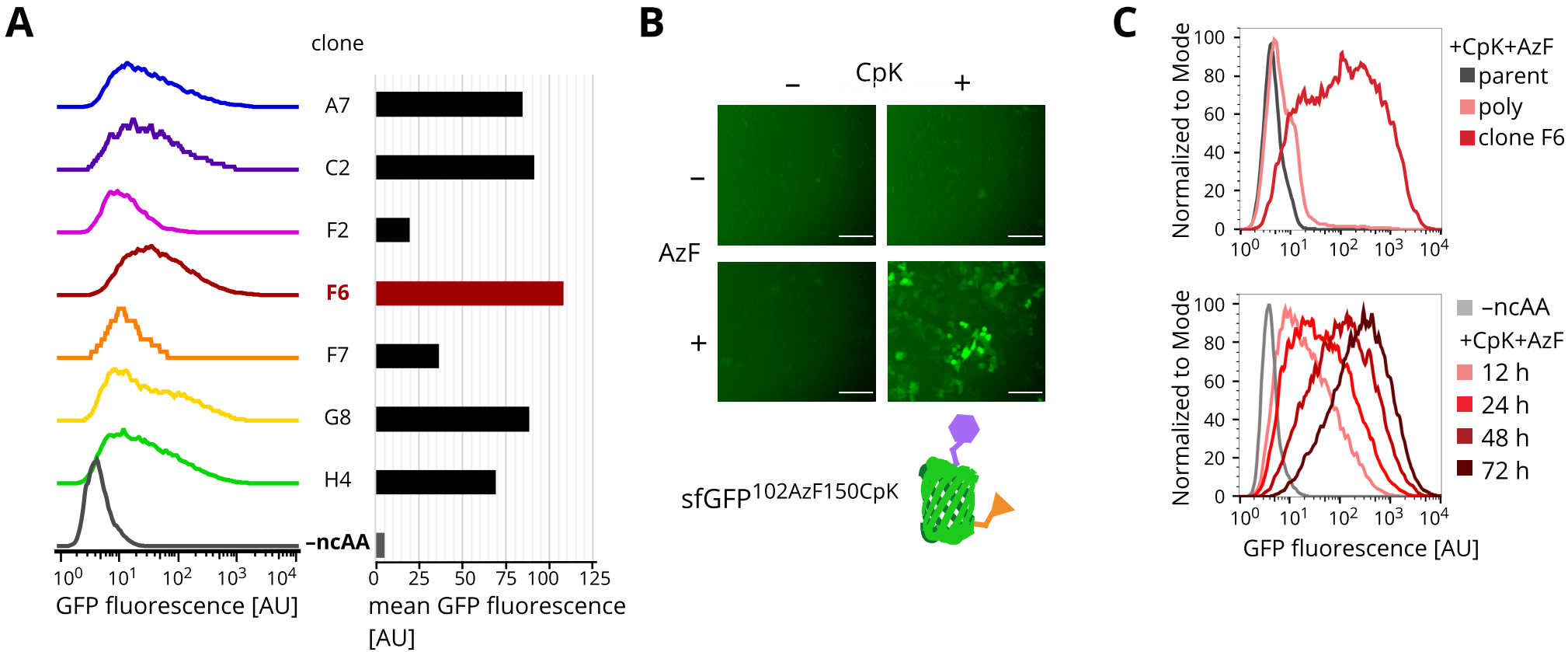
to Figure 2


A
pAS backbone
pUC Ori
AmpR
Ins
EF1>
aaRS or reporter
IRES
Res
pA
pAS
4xtRNA/aaRS
4x 7SK TyrT^CUA
AzFRS
FLAG
NeoR
4x 7SK LeuT^CUA
AnapRS
4x tRNA
TAG
sfGFP
BsdR
pAS
4xtRNA/GFP^150TAG
B
–ncAA
+AzF
clone 1
clone 3
Normalized to Mode
GFP fluorescence [AU]
C
Q3
0.42
Q2
94.6
Q1
2.57
Q1
2.45
Anap fluorescence [AU]
GFP fluorescence [AU]
HCT116
–ncAA
+Anap 2h chase
+Anap no chase
LeuT/AnapRS+GFP^150TAG
–ncAA
+Anap 2h chase
D
Normalized to Mode
GFP fluorescence [AU]
Anap fluorescence [AU]
HCT116
–ncAA 2h chase
+Anap no chase
+Anap 2h chase
LeuT/AnapRS+GFP^150TAG
–ncAA 2h chase
+Anap 2h chase
E
Anap fluorescence [AU]
GFP fluorescence [AU]
–ncAA
+Anap LeuT/AnapRS
+Anap LeuT/AnapRS +GFP^150TAG

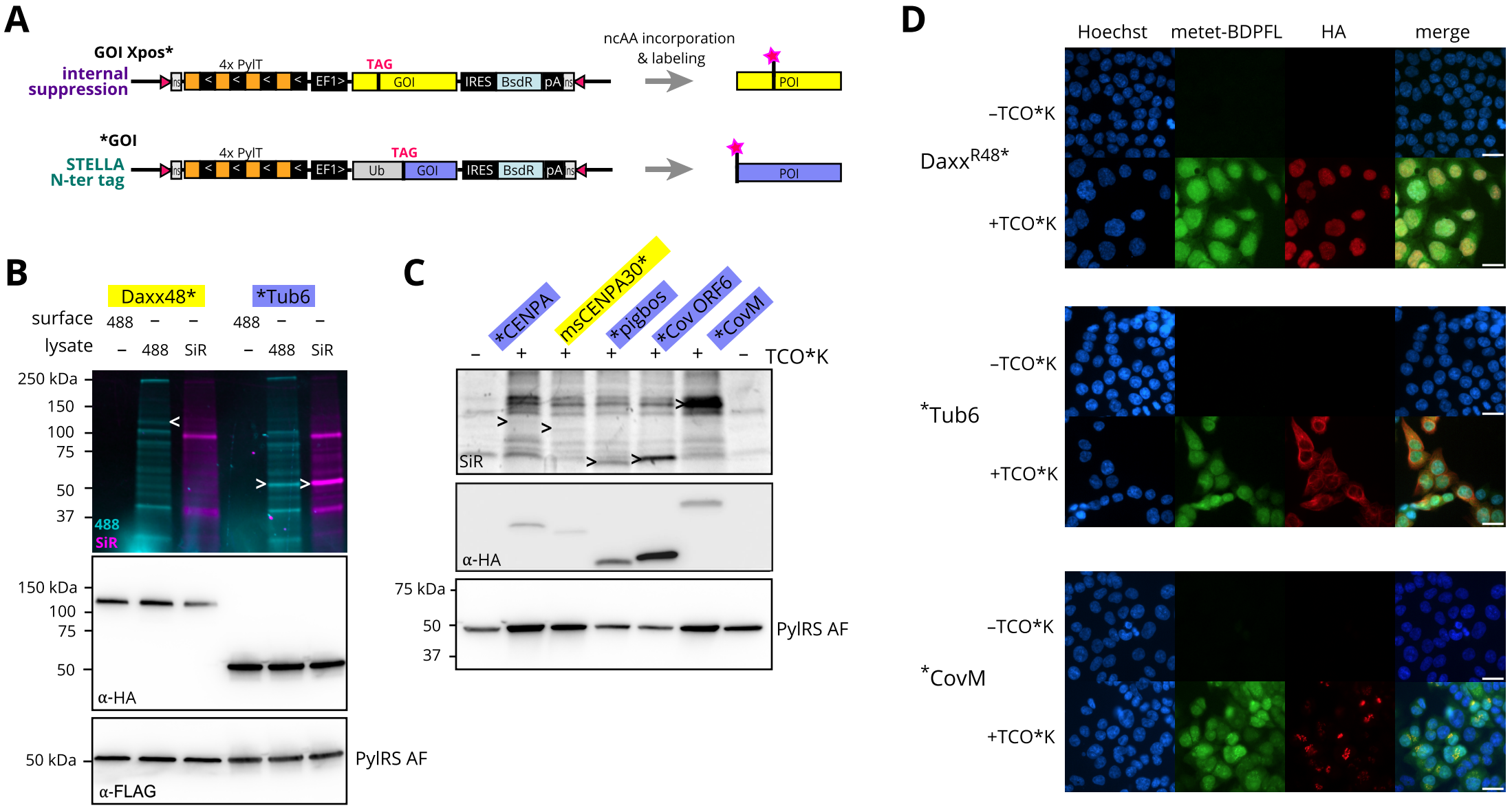
to Figure 3


A
mean GFP fluorescence
median GFP fluorescence
GFP fluorescence [AU]
5000
4000
3000
2000
1000
0
CpK
– + – + – + – + – + – + – + – + – +
tfx
poly
clone D9
clone E11
clone D10
GFP150*
TAG
TAA
TGA
B
GFP150TAG
clone
A5
D9
E7
G2
G6
H6
+CpK
–ncAA
bulk sort
GFP fluorescence [AU]
mean GFP fluorescence [AU]
GFP150TAA
A4
B10
C5
C9
D3
E10
E11
F8
H7
GFP150TGA
D10

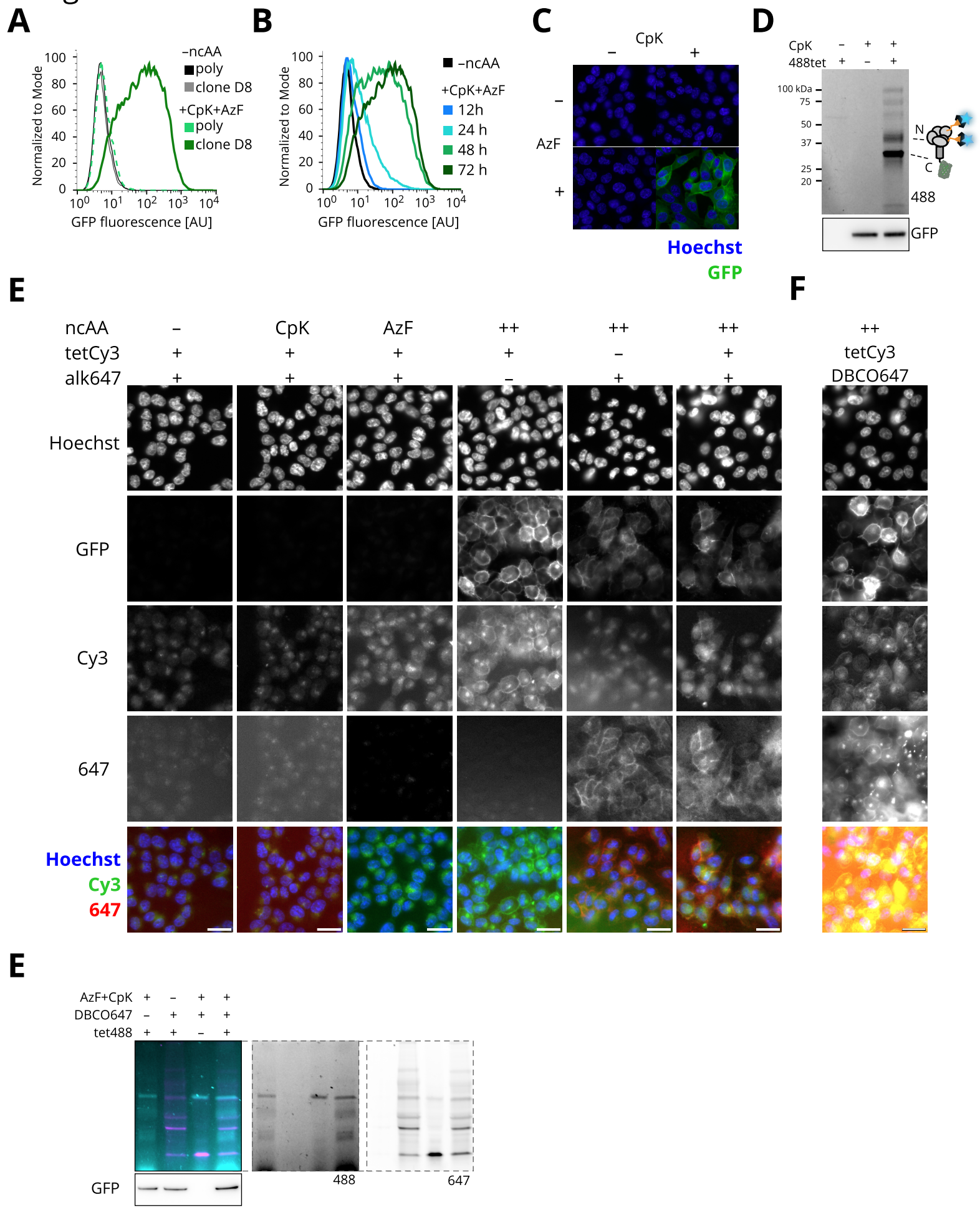
to Figure 4


A
clone
A7
C2
F2
F6
F7
G8
H4
–ncAA
GFP fluorescence [AU]
mean GFP fluorescence [AU]
B
CpK
–
+
AzF
–
+
sfGFP102AzF150CpK
C
Normalized to Mode
+CpK+AzF
parent
poly
clone F6
–ncAA
+CpK+AzF
12 h
24 h
48 h
72 h
GFP fluorescence [AU]

to Figure 5


A
GOI Xpos*
internal suppression
4x PylT
TAG
EF1>
GOI
IRES
BsdR
pA
ncAA incorporation & labeling
POI
*GOI
STELLA N-ter tag
Ub
B
Daxx48*
*Tub6
surface
lysate
488
SiR
250 kDa
150
100
75
50
37
α-HA
150 kDa
α-FLAG
50 kDa
PylRS AF
C
*CENPA
msCENPA30*
*pigbos
*Cov ORF6
*CovM
TCO*K
75 kDa
D
Hoechst
metet-BDPFL
HA
merge
–TCO*K
+TCO*K
Daxx^{R48*}
*Tub6
*CovM

to Figure 6


A
Normalized to Mode
GFP fluorescence [AU]
−ncAA
poly
clone D8
+CpK+AzF
poly
clone D8
B
Normalized to Mode
GFP fluorescence [AU]
−ncAA
+CpK+AzF
12h
24 h
48 h
72 h
C
CpK
AzF
Hoechst
GFP
D
CpK
488tet
100 kDa
75
50
37
25
20
N
C
488
GFP
E
ncAA
tetCy3
alk647
CpK
AzF
Hoechst
GFP
Cy3
647
Hoechst
Cy3
647
F
tetCy3
DBCO647
E
AzF+CpK
DBCO647
tet488
GFP
488
647

Supplementary Figures

Figure S1, relates to Figure 1

A) Quantification of GFP fluorescence in lysates of low selected (2 µg/ml puromycin, 500 µg/ml blasticidin) stable integrant HEK293T cell line. Where indicated (yellow shading), the cell line was transiently transfected with different pAS plasmids bearing 4xPylT and/or PylRS variants. For each condition, quadruplicate transfections were performed. For three of the four samples medium was supplemented with 0.2 mM CpK, all samples were harvested 24 h post transfection. GFP fluorescence signal is normalized to that of untransfected cells cultured in presence of 0.2 mM CpK for 24 h, background fluorescence (–CpK) values were subtracted. Error bars indicate standard deviation. Plasmids transfected by lanes: 1 no transfection, 2 control plasmid $GFP^{102TAA133TGA150TAG}$, 3 PylRS (Addgene #154762), 4 4xPylT (Addgene #140008), 5 4xPylT/RS (Addgene #140009), 6 8xPylT/RS (Addgene #140008), 7 4xM15T/PylRS (Addgene #174890), 8 4xA2-1T/PylRS (A2-1 (Kelemen et al. 2022)), 9 4xPylT/IPYE_*Mba/Mma*chPylRS ((Bryson et al. 2017)).
B) Quantification of mean GFP fluorescence in flow cytometry samples shown in Figure 1D.
C) Immunoblotting of amber suppression with PylRS variant AbKRS-chIPYE (AbKRS). The chemical structure of *N6*-[[2-(3-Methyl-3*H*-diazirin-3-yl)ethoxy]carbonyl]-L-lysine (AbK) is shown. HCT116 cells transiently transfected with PylT/AbKRS and PylT/GFP^{150TAG} plasmids (transient, left), or polyclonal stable integrants with the same plasmids (integration, right) were grown without ncAA (–AbK) or presence of 0.5 mM AbK for 24 h. Immunostaining for GFP expression levels, FLAG-AbKRS and β-actin.
D) Flow cytometry of HCT116 cells transiently transfected with PylT/AbKRS and PylT/GFP^{150TAG} plasmids (transient, top), or polyclonal stable integrants with the same plasmids (integration, bottom) were grown in absence of ncAA (black) or in presence of 0.5 mM AbK (green) for 24 h.

Figure S2, relates to Figure 2

A) Schematic representation of pAS constructs for AzF and Anap incorporation. AzFRS or AnapRS are expressed under EF1α promoter control. An IRES (internal ribosome entry site) followed by a neomycin resistance (NeoR) is positioned immediately downstream of the aaRS. A cassette of four tandem copies of orthogonal tRNA (*B. stearothermophilus* $TyrT^{CUA}$ ($TyrT^{CUA}$) or *E. coli* $LeuT^{CUA}$ ($LeuT^{CUA}$) for 7SK PolIII driven expression is integrated upstream

of the EF1α promoter. We chose to add external 7SK PolIII promoters leading each tRNA, even for *Bst*TyrT^{CUA}, which contains an internal B-box for PolIII transcription (Liu et al. 2007).

B) The entire 4xtRNA/aaRS cassette is flanked by insulator sequences (Ins) and 3' and 5' inverted repeats for genomic integration by PBT (pink triangle). Plasmids for GFP^{150TAG} amber suppression reporter with either TyrT^{CUA} or LeuT^{CUA} cassettes have the same architecture, but contain a blasticidin resistance gene (Bsd).

C) Flow cytometry of amber suppression clonal cell lines isolated from the AzFRS, cognate TyrT^{CUA} tRNA and GFP^{150TAG} reporter integrant polyclonal population shown in Figure 2. GFP fluorescence from the GFP^{150TAG} reporter was measured in the two clonal cell lines after 24 h without ncAA (dashed black line) or in presence of 0.5 mM AzF (green line).

D) Flow cytometry of amber suppression cell line generated with AnapRS, cognate LeuT^{CUA} tRNA and GFP^{150TAG} reporter compared to the parental HCT116 cells. Cells were grown in presence of 0.05 mM Anap for 24 h and either measured directly or after a 2 h chase with ncAA-free medium. The 2D plot for GFP and Anap fluorescence shows the wild type HCT116 populations without Anap (-ncAA, red), with Anap (+Anap, grey) and with Anap washed out (+Anap, chase, black). The LeuT^{CUA}/AnapRS+GFP^{150TAG} integrant stable cell line is shown in blue (-ncAA) and green (+Anap).

E) Same data as in C) presented as histogram plots separating GFP and Anap fluorescences.

F) Flow cytometry of amber suppression cell line generated with AnapRS, cognate LeuT^{CUA} tRNA and a GFP^{150TAG} reporter (green) compared to an integrant cell line with LeuT^{CUA}/AnapRS only (no reporter, blue). Cells were grown in presence of Anap for 24 h and measured after a 4 h chase with ncAA-free medium. The "no reporter" cell line retains Anap.

Figure S3, relates to Figure 3

A) Quantification of mean and median GFP fluorescence measured in Figure 3A.

B) Comparison of GFP fluorescence generated by amber, ochre or opal suppression in the dual suppression clones sorted from polyclonal pools of PylT/PylRS and GFP^{150*} reporter HCT116 integrants generated by PB transgenesis. The clonal populations were grown in presence of 0.2 mM CpK for 24 h. The clones used in Figure 3A are highlighted in red. For comparison, shown in grey, a population grown from 100 cells is also shown as –ncAA control and approximation of the polyclonal pool.

Figure S5, relates to Figure 5

A) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression clones sorted from a common polyclonal pool of PylT^{UUA}/PylRS ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and GFP^{102TAG150TAA} reporter integrants generated by PB transgenesis. The clonal populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 48 h.

B) Fluorescence microscopy images of dual suppressor clonal F6 population grown without ncAA supplement, with either CpK or AzF or both for 48 h. White scale bars indicate 250 µm.

C) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression cell polyclonal pool of PylT^{UUA}/PylRS ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and GFP^{102TAG150TAA} reporter, the derived F6 clonal population or HCT116 parental cells. The populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 48 h before measuring GFP fluorescence by flow cytometry.

D) Time course of ncAA incorporation by the dual suppressor F6 clonal population, cells were grown without ncAAs or with 0.2 mM CpK and 0.5 mM AzF for 12 h, 24 h, 48 h or 72 h prior to measuring GFP fluorescence by flow cytometry.

Figure S5, relates to Figure 5

A) Scheme illustrating construct design and ncAA labeling approaches for internal TAG and STELLA amber suppression. An asterisk (*) is used to indicate position of the amber codon and ncAA incorporation site. The amber stop codon signaling for ncAA incorporation can either be placed within the coding sequence of a gene of interest (GOI), producing the protein of interest (POI) in an ncAA-dependent manner, alternatively the POI can be expressed with an N-terminal ubiquitin-tag separated by an amber codon, leading to POI with ncAA at the N-terminus after ubiquitin processing (STELLA N-terminal tag) (Lafranchi et al. 2020).

B) Amber suppression PB integrant cell lines with PylT/RS AF and either PylT/Daxx48TAG-HA (Daxx48*) or PylT/STELLA N-terminally TAGed Tub6-HA (*Tub6). Stable integrant populations recovered after 7 d selection were grown in presence of 0.1 mM TCO*K for 24 h and SPIEDAC-labeled on live cells with 4 µM tet488 or in lysate with 1 µM tet488 or 1 µM tetSiR. Lysate aliquots were separated by SDS-PAGE and fluorescence captured in-gel prior to membrane transfer and immunoblotting for the C-terminal HA-tag and the FLAG-PylRS AF.

C) In-gel fluorescence and immunoblotting for further PB integrant amber suppression cells lines generated in HCT116 cells with PylT/RS AF and STELLA N-terminally TAGed CENPA, pigbos, *Cov*ORF6 and *Cov*M or murine CENPA Q30TAG (msCENPA30*). Stable integrant populations recovered after 7 d selection were grown in presence of 0.1 mM TCO*K for 24 h and SPIEDAC-labeled in lysate with 1 μM tetSiR. Lysate aliquots were separated by SDS-PAGE and in-gel fluorescence visualized prior to membrane transfer and western blot for the C-terminal HA-tag and the FLAG-PylRS AF.

D) Fluorescence microscopy of amber suppressor HCT116 polyclonal pools encoding PylT/RS AF and a protein of interest: Daxx^{48TAG} or STELLA N-terminally tagged Tub6 or CovM. Cells were grown in presence of 0.1 mM TCO*K for 48 h. Cells were fixed with 4% PFA, permeabilized with 0.1% triton, blocked with 2% BSA and stained with 0.5 μM metetBDPFL, primary anti-HA and secondary 555-coupled antibodies. Prior to imaging on a Nikon Ti2, nuclei were counterstained with Hoechst 33342. White scale bars indicate 25 μm.

Figure S6, relates to Figure 6

A) Comparison of GFP fluorescence generated by dual amber and ochre suppression in the dual suppression cell polyclonal pool of M15^{UUA}/PylRS ochre suppressor, TyrT^{CUA}/AzFRS amber suppressor pair and SynNotch^{204TAG442TAA}-GFP, the derived D8 clonal population or HCT116 parental cells. The populations were grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h before measuring GFP fluorescence by flow cytometry.

B) Time course of ncAA incorporation by the dual suppressor D8 clonal population, cells were grown without ncAAs or with 0.2 mM CpK and 0.5 mM AzF for 12 h, 24 h, 48 h or 72 h prior to measuring GFP fluorescence by flow cytometry.

C) Fluorescence microscopy images of dual suppressor SynNotch^{204TAG442TAA}-GFP clonal D8 population grown without ncAA, with either CpK or AzF or both ncAAs for 48 h. GFP fluorescene (green) and Hoechst 33342 nuclear staining (blue) are shown.

D) Transient transfection of HEK293T cells with PylT/RS and M15^{UUA}/SynNotch^{204TAA442TAG}-GFP plasmids (at 1:4 ratio) grown with and without 0.2 mM CpK. Live cells were surface labeled with 2 μM tet488 48 h post transfection. Aliquots of cell lysates were separated by SDS-PAGE, exposed at 460 nm (488) for in-gel fluorescence and immunostained for GFP and FLAG-RS after membrane transfer.

E) Live-cell imaging of SPIEDAC- and CuAAC-labeled HCT116-derived dual suppressor clonal population with M15^{UUA}/RS, TyrT^{CUA}/AzFRS, and M15^{UUA}/SynNotch^{204TAA442TAG}-GFP integrated by PB transposition. Cells were grown for 72 h in the absence (–ncAA) or the presence of

0.2 mM CpK and 0.5 mM AzF. Surface labeling by SPIEDAC with 2 μM tetCy3 and CuAAC with 2 μM alk647 was carried out prior to Hoechst 33342 counterstaining, fixation with PFA and imaging. White scale bars indicate 25 μm.

F) Conditions like E) but instead of CuAAC with alk647, AzF was labeled by SPAAC with 2 μM DBCO647. White scale bars indicate 25 μm.

G) SPIEDAC and SPAAC labeling on live HCT116 clonal SynNotch^{204CpK442AzF}-GFP cells. grown in presence of 0.2 mM CpK and 0.5 mM AzF for 72 h. Cells were labeled with 2 μM tet488 and 2 μM DBCO647. Lysates were separated by SDS-PAGE, in-gel fluorescence at 460 nm (488) and 630 nm (647) and immunoblotting against the C-terminal GFP are shown.

Supplementary material and method

Noncanonical amino acids

N6-[[2-(3-Methyl-3*H*-diazirin-3-yl)ethoxy]carbonyl]-L-lysine (AbK, CAS: 1253643-88-7, SiChem, SC-8034) was prepared as described for the other ncAAs.

Intracellular SPIEDAC labeling and immunofluorescence microscopy

HCT116 PB integrant polyclonal pools with PylT/PylRS AF and PylT/gene of interest (GOI) were seeded at 15 000 cells per well on poly-L-lysine coated 96 well imaging plates (Ibidi) in presence of 0.1 mM TCO*K for 48 h. Subsequently, the cells were washed with PBS, fixed in 4% PFA) for 10 min, permeabilized in PBS with 0.1% triton X-100 (v/v) (PBS-T), blocked in 2% BSA in PBS-T and stained with 0.5 µM methyltetrazine-BDP-FL (metetBDPFL) (Jena Bioscience) for 1 h at RT, incubated with primary anti-HA antibody (Santa Cruz, sc-7392) followed by incubation with secondary 555-coupled antibody (Life Technologies, A-31570) and 2 µM Hoechst 33342 (Life Technologies). The cells were washed and imaged in PBS on a Nikon Eclipse Ti2 inverted widefield microscope, using a 20x (0.75 NA) objective and filters sets for DAPI, AF488, Cy3 and Cy5 fluorescence.

Supplementary References

Bryson DI, Fan C, Guo L-T, Miller C, Söll D, Liu DR. 2017. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat Chem Biol* **13**: 1253–1260.

Kelemen RE, Jewel D, Huang RL, Zhu Z, Cao X, Pasha M, Anthony J, van Opijnen T, Chatterjee A. 2022. Virus-assisted directed evolution of enhanced suppressor tRNAs in mammalian cells. *BioRxiv*. doi: 10.1101/2022.01.21.477302.

Lafranchi L, Schlesinger D, Kimler KJ, Elsässer SJ. 2020. Universal Single-Residue Terminal Labels for Fluorescent Live Cell Imaging of Microproteins. *J Am Chem Soc* **142**: 20080–20087.

Liu W, Brock A, Chen S, Chen S, Schultz PG. 2007. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* **4**: 239–244.