nature chemistry

SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.1919

Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET

Kaihang Wang[1,2], Amit Sachdeva[1,2], Daniel J. Cox[2], Nabil W. Wilf[2], Kathrin Lang[2], Stephen Wallace[2], Ryan A. Mehl[3] and Jason W. Chin[2*]

[2]Medical Research Council Laboratory of Molecular Biology,

Francis Crick Avenue, Cambridge CB2 0QH, UK

[3]Oregon State University, Department of Biochemistry and Biophysics,

2011 ALS, Corvallis, OR97331

[1]Equal contributors

[*]Correspondence: chin@mrc-lmb.cam.ac.uk

Table of Contents

Plasmids construction

To construct the transplant Pyl $tRNA_{XXXX}$, the anticodon of the Pyl $tRNA_{CUA}$ in the pCDF PylRS/$tRNA_{CUA}$ plasmid (a plasmid expressing Methanosarcina Barkeri (*Mb)* PylRS and *Mb*$tRNA_{CUA}$ from constitutive promoters[1]) was changed by enzymatic inverse PCR using primers PylTtXXXXf and PylTtXXXXr (**Supplementary Table 1**) to make the pCDF PylRS/tRNA(transplant)$_{XXXX}$. Two additional copies of the *Mb*PylRS gene were amplified by PCR using primers pCDF1BamHIf and pCDF1743BglIIr (**Supplementary Table 1**) and cloned in the same orientation into the BamHI site of the pCDF PylRS/tRNA(transplant)$_{XXXX}$ to make the pCDF PylRS$_{x3}$/tRNA(transplant)$_{XXXX}$ plasmids. To create the pCDF PylRS$_{x3}$/tRNA(evolved)$_{XXXX}$ plasmids, two additional copies of the *Mb*PylRS gene were amplified by PCR using primers pCDF1BamHIf and pCDF1743BglIIr (**Supplementary Table 1**) and cloned in the same orientation into the BamHI site of the pCDF PylRS/tRNA(evolved)$_{XXXX}$. To create the pCDF NorKRS$_{x3}$/tRNA(evolved)$_{TACT}$ plasmids, NorKRS[2] was created by Quickchange mutagenesis using the *Mm*PylRS gene as a template and primers MmPylS_Y306Gf, MmPylS_Y306Gr, MmPylS_Y384Ff, MmPylS_Y384Fr, MmPylS_I405Rf, MmPylS_I405Rr to introduce the Y306G, Y384F, I405R triple mutations (**Supplementary Table 1**). The internal BamHI site of NorKRS was removed by Quickchange mutagenesis using primers MmPylS_D318dBamHIf and MmPylS_D318dBamHIr (**Supplementary Table 1**). The NorKRS gene was used to replace the three copies of WT *Mb*PylRS in pCDF PylRS$_{x3}$/tRNA(evolved)$_{TACT}$.

The pO-*gst-cam* plasmid[1] has an orthogonal ribosomal binding site 5' to the ORF of a fusion gene of *glutathione-S-transferase* and *calmodulin* (*gst-cam*) and the GST-CaM

fusion protein is translated by orthogonal ribosome. This plasmid was used as the starting point to construct the O-*gst-cam*$_{1XXXX}$ gene. The 1st ATG Met codon of *calmodulin* (*cam*) was mutated to quadruplet codon (XXXX) or amber codon (TAG) by Quickchange mutagenesis using primers CaM1XXXXf and CaM1XXXXr (**Supplementary Table 1**). In a second round of mutagenesis, the second quadruplet codon (XXXX) or amber codon (TAG) was introduced at either position 40 to replace the CTT Leu codon using primers CaM40XXXXf and CaM40XXXXr or at position 149 to replace the AAG Lys codon using primers CaM149XXXXf and CaM149XXXXr (**Supplementary Table 1**). Please note the different numbering in previous paper[1], which is based on the numbering in PDB structure 4CLN[3]. The 40th amino acid Gly in the previous paper[1] corresponds to the 41st GGA Gly codon in this work. The O-*gst-cam*$_{1TAGA}$ gene was amplified by PCR using primers GstCaMgeneF_KpnI and GstCaMgeneR_KpnI (**Supplementary Table 1**) and cloned into the KpnI site of the pRSF ribo-Q1 (expressing ribo-Q1 rRNA from an IPTG inducible promoter[1]) to make the pRSF ribo-Q1 O-*gst-cam*$_{1TAGA}$. The complete set of pRSF ribo-Q1 O-*gst-cam*$_{1XXXX}$ and pRSF ribo-Q1 O-*gst-cam*$_{1XXXX+40/149XXXX}$ plasmids were created by cloning the O-*gst-cam*$_{1XXXX}$ and O-*gst-cam*$_{1XXXX+40/149XXXX}$ genes into pRSF ribo-Q1 O-*gst-cam*$_{1TAGA}$, replacing *cam*$_{1TAGA}$ between SwaI and SpeI restriction sites. Constructs for translation by the wild-type ribosome were created using p-*gst-cam* plasmid as a template.

pDULE *Mj*AzPheRS*/tRNA$_{UCCU}$ is previously described[1]. It contains an *Mj*TyrRS variant that recognises **4** and aminoacylates *Mj*tRNA$_{UCCU}$. The pSUP *Mj*AzPheRS/tRNA$_{CUA}$[4] expresses *Mj*AzPheRS, which recognizes **4**, and expresses multiple copies of *Mj*tRNA$_{CUA}$ from constitutive promoters. The *Mj*AzPheRS in pSUP *Mj*AzPheRS/tRNA$_{CUA}$ is replaced with TetPheRS[5], PrpTyrRS[6] or BpaRS[7] to

make pSUP *Mj*TetPheRS/tRNA$_{CUA}$, pSUP *Mj*PrpTyrRS/tRNA$_{CUA}$, and pSUP *Mj*BpaRS/tRNA$_{CUA}$.

Construction of Pyl tRNA(N8)$_{XXXX}$ and quadruplet decoding reporters

To select evolved Pyl tRNA$_{XXXX}$ that efficiently decode distinct quadruplet codons TAGA, AGGA, AGTA and CTAG, we created saturation mutagenesis libraries in the anticodon stem loop region of each of the four Pyl tRNA(N8)$_{XXXX}$ libraries. Nucleotides (nt 30-33 and 37-40) were randomized for Pyl tRNA(N8)$_{UCUA}$. Nucleotides 29-32 and 38-41 were randomized for Pyl tRNA(N8)$_{UCCU}$, Pyl tRNA(N8)$_{UACU}$ and Pyl tRNA(N8)$_{CUAG}$ (**Figure-1b**). These libraries were created by enzymatic inverse PCR using pCDF PylRS/tRNA$_{CUA}$ as template and library construction primers PylTN8XXXXf and PylTN8XXXXr listed in **Supplementary Table 1**. Each library has a diversity of greater than 10^8, ensuring more than 99.9% coverage.

To create a positive selection marker for quadruplet decoding tRNAs that are efficiently read by ribo-Q1, we used a previously described O-*cat* plasmid as a template[1]. The 111th triplet codon GAC (encoding Asp) in the *cat* ORF was replaced by a quadruplet codon (TAGA/AGGA/AGTA/CTAG) by Quickchange mutagenesis using primers catD111XXXXf and catD111XXXXr to make the positive selection marker O-*cat*$_{111XXXX}$ (**Supplementary Table 1)**.

To create a negative selection marker for quadruplet decoding tRNAs that are efficiently read by ribo-Q1, we first replaced the ORF of *cat* in the O-*cat* plasmid with the ORF of *barnase*$_{3UAG+45UAG}$, in which the toxic ribonuclease Barnase contains UAG codons at position 3 and 45 of its gene[8], by PCR using primers G9_BarnaseF_NotI and BarnaseR_PmeI (**Supplementary Table 1**) and cloning

between NotI and PmeI restriction sites to make the *O-barnase*$_{3UAG+45UAG}$ plasmid. Quadruplet codons were introduced to replace the UAG stop codons to make the negative selection markers *O-barnase*$_{3XXXX+45XXXX}$ by PCR with listed primer pairs G9_BarAGGAx2F and G9_BarAGGAx2R (**Supplementary Table 1)** and the PCR products were cloned between NotI and XmaI restriction sites.

Selection of Pyl tRNA(N8)$_{XXXX}$ for enhanced quadruplet decoding in the ribo-Q1 orthogonal mRNA system

We transformed each Pyl tRNA(N8)$_{XXXX}$ library into GeneHogs *E. coli* cells bearing the matching *O-barnase*$_{3XXXX+45XXXX}$ plasmid and pRSF ribo-Q1 plasmid. The transformed cells were recovered in 1 ml SOB medium containing 2% glucose for one hour at 37°C. The transformation was used to inoculate 200 ml of LB-GKST (LB media with 2% glucose, 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 6.25 μg ml^{-1} tetracycline) and incubated overnight (37°C, 250 rpm, 16 h). 2 ml of the overnight culture (OD_{600}≈2) was pelleted by centrifugation (3,000g) and washed with LB to remove glucose, prior to diluting to OD_{600}=0.1 in LB-KST (LB media with 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 6.25 μg ml^{-1} tetracycline), and incubated (37°C, 250 rpm, 2-3 h) until OD_{600} reached 0.5. IPTG (1mM final concentration) was added to induce the expression of ribo-Q1 rRNA at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h, when the OD_{600} reached approximately 2. Aliquots (500 μl, OD_{600}≈2) of the culture were serial diluted and plated on LB-KST agar plate supplemented with 1mM IPTG. The plates were incubated at 37°C for 16 hours. Surviving colonies were scraped from the plate and plasmid DNA was extracted by Qiagen Miniprep Kit to make the post negative selection pool.

The negative selection was followed by a positive selection. The post negative selection Pyl tRNA(N8)$_{XXXX}$ pools were cleansed of non pCDF PylRS/tRNA$_{XXXX}$ plasmids by digestion with SacII to remove the *O-barnase*$_{3XXXX+45XXXX}$ and pRSF ribo-Q1 plasmid. The cleansed post negative selection Pyl tRNA(N8)$_{XXXX}$ pools were electroporated into GeneHog cells holding the matching O-*cat*$_{111XXXX}$ and pRSF ribo-Q1 plasmids. The transformed cells were recovered in 1 ml SOB medium containing 2% glucose for one hour at 37°C. The transformation was used to inoculate 200 ml of LB-GKST (LB media with 2% glucose, 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 6.25 μg ml^{-1} tetracycline) and incubated overnight (37°C, 250 rpm, 16 h). 2ml of the overnight culture (OD_{600}≈2) was pelleted by centrifugation (3,000g) and washed with LB to remove glucose, prior to diluting to OD_{600}=0.1 in LB-KST (LB media with 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 6.25 μg ml^{-1} tetracycline), and incubated (37°C, 250 rpm, 2-3 h) until OD_{600} reached 0.5. IPTG (1 mM final concentration) and **1** (8 mM final concentration) were added, to induce the expression of ribo-Q1 rRNA and facilitate aminoacylation of Pyl tRNA(N8)$_{XXXX}$ library members by PylRS at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h, when the OD_{600} reached approximately 2. Aliquots (500 μl, OD_{600}≈2) of the culture were serial diluted and plated on LB-KST agar plates supplemented with 1 mM IPTG, 8 mM **1**, and chloramphenicol of different concentrations (100 μg ml^{-1}, 150 μg ml^{-1}, 200 μg ml^{-1}). The plates were incubated at 37°C for 40 hours.

To separate the selected pCDF PylRS/tRNA(N8)$_{XXXX}$ plasmid from the O-*cat*$_{111XXXX}$ and pRSF ribo-Q1 plasmids, total plasmid DNA from selected clones was purified, digested with SacII, transformed into DH10B cells and plated on LB agar

supplemented with spectinomycin (75 μg ml^{-1}). Individual colonies from the spectinomycin plates were incubated overnight in LB media supplemented with spectinomycin (75 μg ml^{-1}) and replica plated onto three sets of LB agar plates supplemented with spectinomycin (75 μg ml^{-1}), or tetracycline (25 μg ml^{-1}), or kanamycin (50 μg ml^{-1}). Cells containing pCDF PylRS/tRNA(N8)$_{XXXX}$ only grew on spectinomycin plates but not tetracycline or kanamycin plates. The DNA from separated pCDF PylRS/tRNA(N8)$_{XXXX}$ clones were purified, confirmed by restriction digestion and DNA sequencing.

Characterization of evolved Pyl tRNA(N8)$_{XXXX}$ with enhanced quadruplet decoding via a chloramphenicol reporter assay

To quantify the quadruplet decoding activity of selected Pyl tRNA(N8)$_{XXXX}$ clones, the selected pCDF PylRS/tRNA(N8)$_{XXXX}$ plasmids were cotransformed with O-*cat*$_{111XXXX}$ and pRSF ribo-Q1. Cells were recovered in 1 ml of SOB media for 1 h, inoculated into 10 ml of LB-KST and incubated overnight (37°C, 250 rpm, 16 h). 1 ml of the resulting culture was inoculated to 9 ml of fresh LB-KST and incubated for 90 min (37°C, 250 rpm). 1 ml of the 90 min culture was inoculated into 9 ml of LB-KST supplemented with 1.1 mM IPTG only to make the "-ve" fraction, another 1 ml of the same culture was inoculated to 9 ml of LB-KST supplemented with 1.1 mM IPTG plus 8.8 mM **1** to make the "+ve" fraction. Both fractions were incubated in parallel for 4 h (37°C, 250 rpm). Cultures (1 ml) of individual clones' "+ve" fractions were transferred to a 96-well block and replica arrayed in a 96-well format onto LB-KST agar plates supplemented with 8 mM **1**, 1mM IPTG and chloramphenicol of different concentrations from 0 to 500 μg ml^{-1}. Cultures (1 ml) of individual clones' "-ve" fractions were replica arrayed in the same way onto LB-KST agar plates

supplemented with 1mM IPTG only and chloramphenicol of different concentrations from 0 to 500 μg ml^{-1}. The clones giving the highest level of chloramphenicol resistance in the presence of **1** and minimal chloramphenicol resistance in the absence of **1** were further characterized.

Single incorporation and double incorporation of unnatural amino acids into recombinant proteins

For single incorporation, MDS42 recA or Genehogs *E. coli* cells were transformed by electroporation with pRSF ribo-Q1 O-*gst-cam*$_{1XXXX}$ (or pRSF ribo-Q1 O-*gst-cam*$_{40XXXX}$) and pCDF PylRS$_{.3}$/tRNA(evolved or transplanted)$_{XXXX}$, and recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KS (LB media with 25 μg ml^{-1} kanamycin, and 37.5 μg ml^{-1} spectinomycin) and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KS and incubated (37°C, 250 rpm) until OD_{600} reached 0.5. IPTG (1 mM final concentration) without or with unnatural amino acids (**1**, **2**, **3** were used at 8 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS. For incorporation of **4** into O-*gst-cam*$_{1AGGA}$, MDS42 recA *E. coli* cells were transformed by electroporation with pRSF ribo-Q1 O-*gst-cam*$_{1AGGA}$ and pDULE *Mj*AzPheRS*/$tRNA_{TCCT}$, and recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KT (LB media with 25 μg ml^{-1} kanamycin, and 12.5 μg ml^{-1} tetracycline) and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KT and incubated (37°C, 250 rpm). IPTG (1 mM final concentration) without or with **4** (5 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250

rpm) for 4 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS. For incorporation of **1** into *gst-cam*$_{1XXXX}$ using the evolved Pyl tRNA$_{XXXX}$ and endogenous wild type ribosomes, MDS42 recA *E. coli* cells were transformed by electroporation with RSF wt-Ribo *gst-cam*$_{1XXXX}$ and the pCDF PylRS$_{.3}$/ evolved tRNA$_{XXXX}$, and recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KS and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KS and incubated (37°C, 250 rpm). IPTG (1mM final concentration) without or with **1** (8 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS. Other anticodon, codon pair experiments were performed in an analogous manner.

For incorporation of **1** in response to UAG and **4** in response to AGGA using a published four plasmids system[1], MDS42 recA *E. coli* cells bearing pRSF ribo-Q1 plasmid and O-*gst-cam*$_{1AGGA-40TAG}$ (numbering from the 1st ATG Met codon of CaM, the 1st ATG Met codon is changed to AGGA, and the 40th CTT Leu codon is changed to TAG. Please note the difference in numbering in the previous paper[1], which is based on the numbering in PDB structure 4CLN. The 40th amino acid Gly[1] corresponds to the 41st GGA Gly codon in this work) were co-transformed by electroporation with pDULE AzPheRS*/tRNA$_{UCCU}$ and pCDF PylRS/tRNA$_{CUA}$, and recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KAST (LB media with 12.5 μg ml^{-1} kanamycin, 25 μg ml^{-1} ampicillin, 18.8 μg ml^{-1} spectinomycin, and 6.3 μg ml^{-1} tetracycline) and incubated overnight (37°C, 250

rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KAST and incubated (37°C, 250 rpm). IPTG (1 mM final concentration) without or with unnatural amino acids (**1** was used at 8 mM, and **4** at 5 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 6 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS.

For incorporation of **1** in response to UAG and **4** in response to AGGA using an improved three plasmids system in which ribo-Q1 rRNA and O-*gst-cam* are encoded on the same plasmid. MDS42 recA *E. coli* cells bearing pRSF ribo-Q1 O-*gst-cam*$_{1AGGA\text{-}40TAG}$ or pRSF ribo-Q1 O-*gst-cam*$_{1TAG\text{-}40AGGA}$ plasmid were co-transformed by electroporation with pDULE AzPheRS*/$tRNA_{UCCU}$ and pCDF $PylRS_{\times 3}$/$tRNA_{CUA}$, and recovered in 1 ml SOB medium for one hour at 37°C. 100 ml LB-KST (LB media with 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 6.3 μg ml^{-1} tetracycline) was inoculated with the recovered transformation and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KST and incubated (37°C, 250 rpm). IPTG (1 mM final concentration) without or with unnatural amino acids (**1** was used at 8 mM, and **4** at 5 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 6 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS. For incorporation of **1** and **4** in response to UAGA and AGGA codons, pRSF ribo-Q1 O-*gst-cam*$_{1AGGA\text{-}40TAG}$ plasmid was replaced with pRSF ribo-Q1 O-*gst-cam*$_{1AGGA\text{-}40TAGA}$, and pCDF $PylRS_{\times 3}$/$tRNA_{CUA}$ was replaced with pCDF $PylRS_{\times 3}$/evolved $tRNA_{UCUA}$.

For incorporation of **1**, **2**, **3** in response to the quadruplet codon AGTA and **4**, **5**, **6**, **7** in response to the UAG codon, MDS42 recA *E. coli* cells bearing pRSF ribo-Q1 O-*gst-cam*$_{1TAG\text{-}40AGTA}$ were co-transformed by electroporation with pCDF PylRS$_{\text{-}3}$/tRNA(evolved)$_{UACU}$ and pSUP AzPheRS/tRNA$_{CUA}$ for **4**, or pSUP *Mj*TetPheRS/tRNA$_{CUA}$ for **5**, or pSUP *Mj*PrpTyrRS/tRNA$_{CUA}$ for **6**, or pSUP *Mj*BpaRS/tRNA$_{CUA}$ for **7**. Transformations were recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KSC (LB media with 12.5 μg ml^{-1} kanamycin, 18.8 μg ml^{-1} spectinomycin, and 8.8 μg ml^{-1} chloramphenicol) and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KSC and incubated (37°C, 250 rpm). IPTG (1 mM final concentration) without or with unnatural amino acids (**1**, **2**, **3** were used at 8 mM, **4**, **6**, **7** were used at 5 mM, **5** at 1 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h. 100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS.

For incorporation of **8** in response to the quadruplet codon AGTA and **5** in response to the UAG codon, MDS42 recA *E. coli* cells bearing pRSF ribo-Q1 O-*gst-cam*$_{1TAG\text{-}40AGTA}$ or pRSF ribo-Q1 O-*gst-cam*$_{1TAG\text{-}149AGTA}$ were co-transformed by electroporation with pCDF NorKRS$_{\text{-}3}$/tRNA(evolved)$_{UACU}$ and pSUP *Mj*TetPheRS/tRNA$_{CUA}$. The transformation was recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 100 ml LB-KSC and incubated overnight (37°C, 250 rpm, 16 h). The overnight culture was diluted to OD_{600}=0.1 in LB-KSC and incubated (37°C, 250 rpm). IPTG (1 mM final concentration) without or with **8** (1 mM) and **5** (1 mM) was added to the culture at OD_{600}≈0.5. The culture was incubated (37°C, 250 rpm) for 4 h.

100 ml of the culture with or without supplemented unnatural amino acids was pelleted (3,000g, 10 min, 4°C) and washed twice with 1 ml PBS.

To purify overexpressed proteins, cell pellets were resuspended in 1 ml of Novagen BugBuster Protein Extraction Reagent (supplemented with 1× Roche protease inhibitor cocktail tablet, 1 mg ml^{-1} Sigma lysozyme, 1 mg ml^{-1} Sigma DNase I) and lysed (25°C, 250 rpm, 1 h). The lysate was clarified by centrifugation (25,000g, 30 min, 4°C). GST containing proteins from the lysate were bound in batch (1 h, 4°C) to 70 μl of glutathione sepharose beads (GE Healthcare). Beads were washed 4 times with 1 ml PBS prior to elution by heating in 1× Invitrogen NuPAGE LDS sample buffer (95°C, 5 min) supplemented with 100 mM DTT. All samples were analysed on 12% Bis-Tris gels (Invitrogen) with BIO-RAD Low Range Molecular Weight Standard as marker. The molar percentage of each band was calculated based on band intensities on Coomassie Blue (InstantBlue, Expedeon) stained gel and molecular weights of full length and truncated proteins.

Purification of CaM u.a.a.$_1$, CaM u.a.a.$_{40}$, CaM u.a.a.$_1$-u.a.a.$_{40}$, and CaM u.a.a.$_1$-u.a.a.$_{149}$

To purify overexpressed proteins for mass spectroscopy, cell pellets were resuspended in 1 ml of Novagen BugBuster Protein Extraction Reagent (supplemented with 1× Roche protease inhibitor cocktail tablet, 1 mg ml^{-1} Sigma lysozyme, 1 mg ml^{-1} Sigma DNase I), washed with PBS, and lysed (25°C, 250 rpm, 1 h). The lysate was clarified by centrifugation (25,000g, 30 min, 4°C). GST-CaM containing proteins from the lysate were bound in batch (1 h, 4°C) to glutathione sepharose beads (GE Healthcare). Beads were washed 4 times with 1 ml PBS prior to adding 300 μl PBS with 2 Units of

thrombin (Invitrogen), which recognizes the thrombin cleavage site between the fused GST and CaM. The CaM was cleaved overnight (16 h, 25°C, 20 rpm). The 300 µl supernatant after the overnight cleavage was transferred to a fresh tube and the glutathione sepharose beads were washed 3 times with 200 µl PBS. All washing fractions were pooled together with the original 300 µl supernatant to reach a total volume of 900µl before adding 100 µl of Ni NTA Agarose beads (Qiagen) and binding for 1 h (4°C, 20 rpm). The cleaved CaM was purified by its C-terminal His×6 tag. Beads were washed 4 times with 1 ml PBS supplemented with 10 mM imidazole and eluted with 4 fractions of 60 µl PBS supplemented with 250 mM imidazole. The four fractions of eluted CaM were pooled together to a total volume of 240 µl. Salt and imidazole was removed using 7K MWCO Zeba spin desalting columns (Thermo Scientific). 15 µl of the purified CaM was mixed with 35 µl of H_2O for electrospray ionization mass spectrometry.

Electrospray ionization mass spectrometry

Mass spectra for protein samples were acquired on an Agilent 1200 LC-MS system that employs a 6130 Quadrupole spectrometer. The solvent system used for liquid chromatography (LC) was 0.2 % formic acid in H_2O as buffer A, and 0.2 % formic acid in acetonitrile (MeCN) as buffer B. Samples were injected into Phenomenex Jupiter C4 column (150 x 2 mm, 5 µm) and subsequently into the mass spectrometer using an fully automated system. Spectra were acquired in the positive mode and analyzed using the MS Chemstation software (Agilent Technologies). The deconvolution program provided in the software was used to obtain the mass spectra. Theoretical average molecular weight of proteins with unnatural amino acids was

calculated by first computing the theoretical molecular weight of wild-type protein using an online tool (http://www.peptidesynthetics.co.uk/tools/), and then manually correcting for the theoretical molecular weight of unnatural amino acids.

Labelling of CaM5$_1$ with 10 and CaM8$_{40}$ with 9 via inverse electron-demand Diels-Alder reactions and determining on-protein labelling reaction rate

To 8 µl of 2.5 µM purified CaM**5**$_1$ in water (≈20 pmol), 1 µl of 2 mM stock of **10** (2 nmole, ≥100 fold excess) in DMF and 1ul of 10xPBS was added. Samples were incubated at 25°C for 15 min to 24 hours before mixed with 4 µl of NuPAGE LDS sample buffer and DTT (100 mM final concentration), heated at 95ºC for 5 min, and loaded onto 12% Bis-Tris gels (Invitrogen) with BIO-RAD Low Range Molecular Weight Standard as marker. The same amount of CaM**8**$_{40}$ was incubated with the same concentration of **10** for 24 hours as control. The labelling resulted in a migration shift on the SDS-PAGE. The fluorescent bands on the gel were imaged and quantified using Typhoon Trio phosphoimager (GE Life Sciences) with excitation filter set to 532 nm and emission filter set to 580 nm. The on-protein labelling reaction rate was calculated following the increase of fluorescence on the labelled protein. The unlabelled CaM**5**$_1$ and the quantitatively labelled CaM**5-10**$_1$ after incubation for 24 hours were purified using 7K MWCO Zeba spin desalting columns (Thermo Scientific) for electrospray ionization mass spectrometry.

To 8 µl of 2.5 µM purified CaM**8**$_{40}$ in water (≈20 pmol), 1 µl of 2 mM stock of **9** (2 nmole, ≥100 fold excess) in DMF and 1ul of 10xPBS was added. Samples were incubated at 25°C for 15 min to 24 hours before mixed with 4 µl of NuPAGE LDS sample buffer and DTT (100 mM final concentration), heated at 95ºC for 5 min, and loaded onto 12% Bis-Tris gels (Invitrogen) with BIO-RAD Low Range Molecular

Weight Standard as marker. The same amount of CaM**5**$_1$ was incubated with the same concentration of **9** for 24 hours as control. The labelling resulted in a migration shift on the SDS-PAGE. The fluorescent bands on the gel were imaged and quantified using Typhoon Trio phosphoimager (GE Life Sciences) with excitation filter set to 488 nm and emission filter set to 526 nm. The on-protein labelling reaction rate was calculated following the increase of fluorescence on the labelled protein. The unlabelled CaM**8**$_{40}$ and the quantitatively labelled CaM**8-9**$_{40}$ after incubation for 24 hours were purified using 7K MWCO Zeba spin desalting columns (Thermo Scientific) for electrospray ionization mass spectrometry.

Labelling of CaM5$_1$-8$_{149}$ and CaM5$_1$-8$_{40}$ with 9 and 10 via mutually orthogonal inverse electron-demand Diels-Alder reactions

To 80 µl of 2.5 µM purified CaM**5**$_1$-**8**$_{149}$ or CaM**5**$_1$-**8**$_{40}$ in water (≈200 pmol), 10 µl of 2 mM stock of **9** (20 nmole, ≥100 fold excess) in DMF and 10ul of 10xPBS was added. The sample was incubated at 25^{o}C for 48 hours to facilitate complete labelling. Excess **9** was removed by running the sample through 7K MWCO Zeba spin desalting columns (Thermo Scientific) (pre-equilibrated with water) three times. The resulting singly labelled CaM**5**$_1$-**8-9**$_{149}$ or CaM**5**$_1$-**8-9**$_{40}$ was concentrated using Amicon Ultra 3K centrifugal filters (Millipore) and analysed by LC-MS or SDS-PAGE. For analysis using SDS-PAGE, samples were mixed with NuPAGE LDS sample buffer supplemented with DTT (100 mM final concentration), heated for 5 min at 95^{o}C and loaded onto 12% Bis-Tris gels (Invitrogen) with BIO-RAD Low Range Molecular Weight Standard as marker. The fluorescent bands on the gel were imaged using Typhoon Trio phosphoimager (GE Life Sciences).

To 80 µl of 2.5 µM purified CaM**5**$_1$-**8-9**$_{149}$ and CaM**5**$_1$-**8-9**$_{40}$ from the previous step or CaM**5**$_1$-**8**$_{149}$ and CaM**5**$_1$-**8**$_{40}$ in water (≈200 pmol), 10 µl of 2 mM stock of **10** (20 nmole, ≥100 fold excess) in DMF and 10ul of 10xPBS was added. The sample was incubated at 25^{o}C for 24 hours to facilitate complete labelling. Excess **10** was removed by running the sample through 7K MWCO Zeba spin desalting columns (Thermo Scientific) (pre-equilibrated with water) three times. The resulting doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$ and CaM**5-10**$_1$-**8-9**$_{40}$ or singly labelled CaM**5-10**$_1$-**8**$_{149}$ and or CaM**5-10**$_1$-**8**$_{40}$ were concentrated using Amicon Ultra 3K centrifugal filters (Millipore) and analysed by electrospray ionization mass spectrometry or SDS-PAGE. The fluorescent bands on the SDS-PAGE gel were imaged using Typhoon Trio phosphoimager (GE Life Sciences).

Fluorescence measurements and Förster resonance energy transfer (FRET) experiments

The fluorescence spectra of CaM**5-10**$_1$-**8-9**$_{149}$ were measured on Luminescence Spectrometer (Perkin Elmer LS55) using singly labelled CaM**5**$_1$-**8-9**$_{149}$ (donor only) and CaM**5-10**$_1$-**8**$_{149}$ (acceptor only) as controls. The fluorescence spectra of CaM**5-10**$_1$-**8-9**$_{40}$ were measured on Luminescence Spectrometer (Perkin Elmer LS55) using singly labelled CaM**5**$_1$-**8-9**$_{40}$ (donor only) and CaM**5-10**$_1$-**8**$_{40}$ (acceptor only) as controls.

In order to measure FRET between the BODIPY-FL and BODIPY-TMRX labelled at position 1 and 149 on CaM**5-10**$_1$-**8-9**$_{149}$, the emission spectra were acquired between 500 nm and 650 nm using an excitation wavelength of 485 nm. The excitation and emission bandwidth were set to 10 nm for all measurements. All fluorescence spectra of doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$ and donor only CaM**5**$_1$-**8-9**$_{149}$ were normalised by

total emission (the total area below the spectra curve). The spectra of acceptor only CaM**5-10**$_1$-**8**$_{149}$ was normalised by matching the acceptor emission at 570 nm by acceptor excitation at 543 nm to that of the doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$. In order to measure FRET between the BODIPY-FL and BODIPY-TMRX labelled at position 1 and 40 on CaM**5-10**$_1$-**8-9**$_{40}$, the emission spectra were acquired between 505 nm and 650 nm using an excitation wavelength of 485 nm. The excitation and emission bandwidth were set to 10 nm for all measurements. All fluorescence spectra of doubly labelled CaM**5-10**$_1$-**8-9**$_{40}$ and donor only CaM**5**$_1$-**8-9**$_{40}$ were normalised by total emission (the total area below the spectra curve). The spectra of acceptor only CaM**5-10**$_1$-**8**$_{40}$ was normalised by matching the acceptor emission at 570 nm by acceptor excitation at 543 nm to that of the doubly labelled CaM**5-10**$_1$-**8-9**$_{40}$.

The effect of urea concentration on the fluorescence spectra of doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$ was measured by diluting 500 pmol of labelled protein with 200 µl of Ca^{2+}-free buffer (final concentration: 30 mM MOPS, 100 mM KCl, 10 mM EDTA, 0.005% Brij35 and 0.1% PEG-8000[9]), and adding urea powder to reach defined concentrations. FRET efficiency (**ε**) was calculated as, $\varepsilon = 1 - I_{DA}/I_D$, where I_{DA} is the donor fluorescence intensity at emission wavelength 515 nm ($I_{515\ nm}$) from doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$, and I_D is the donor fluorescence intensity at emission wavelength 515 nm ($I_{515\ nm}$) from CaM labelled only with the donor (CaM**5**$_1$-**8-9**$_{149}$), i.e. BODIPY-FL.

The effect of Ca^{2+} concentration on the fluorescence spectra of doubly labelled CaM**5-10**$_1$-**8-9**$_{40}$ was measured by diluting 500 pmol of labelled protein with 200 µl of Ca^{2+}-free buffer (final concentration: 30 mM MOPS, 100 mM KCl, 10 mM EDTA, 0.005% Brij35 and 0.1% PEG-8000[9]) and titrating in Ca^{2+} in the same buffer to

defined Ca^{2+} concentrations (10^{-9} to 3×10^{-2} M). The buffer solutions were routinely pre-treated with Chelex 100 resin (Bio-Rad) to remove any possible trace amount of Ca^{2+} by passing 100ml stock 10× buffer (300 mM MOPS, 1 M KCl, 100 mM EDTA, 0.05% Brij35 and 1% PEG-8000) through a column containing 6 g of Chelex 100 in the K^+ form[10]. The untreated 10× buffer were routinely contaminated with 10^{-6} to 10^{-5} M Ca^{2+} as determined using Patton-Reeder indicator and EDTA complexometric titration. The relative donor fluorescence intensity (I') is calculated as $I'=(I_i - I_0)/(I_{MAX} - I_0)$ where I_i is the donor fluorescence intensity at emission wavelength 515 nm from doubly labelled CaM**5-10**$_1$**-8-9**$_{40}$ at a given Ca^{2+} concentration, I_0 is the donor fluorescence intensity at emission wavelength 515 nm at the minimum Ca^{2+} concentration, and I_{MAX} is the donor fluorescence intensity at emission wavelength 515 nm at the maximum Ca^{2+} concentration.

All measurements were repeated at least six times and the error bars represent the standard deviation. The following equation was used for fitting the data: $I'= (a_1[Ca^{2+}]/([Ca^{2+}] + K_1))+ (a_2[Ca^{2+}]/([Ca^{2+}] + K_2))$. Values for a_1, K_1, a_2 and K_2 obtained from fitting are 0.491±0.0314, $(1.125\pm0.3402)\times10^{-6}$ M^{-1}, 0.544±0.0482, and $(1.711\pm0.6590)\times10^{-3}$ M^{-1} respectively. Single-component fitting of the data resulted in a very poor fit with $R^2 = 0.65$ (**Supplementary Fig. 11**). The Förster distance of 60.28 Å for the FRET probes was calculated using the following equation: $R_o= [8.8\times10^{23}. k^2 . n^{-4} . QY_D . J(\lambda)]^{1/6}$ Å, where k^2 (dipole orientation factor) = 2/3 assuming randomly oriented donor and acceptor, n (refractive index)= 1.33 (for water) and QY_D (fluorescence quantum yield of donor in absence of acceptor)= 0.9. J(λ) (the spectral overlap integral) was calculated using the emission spectra of donor (BODIPY-FL), and excitation spectra of the acceptor (BODIPY-TMR-X). The acceptor extinction coefficient (ε_A) of 60,000 was used for this calculation. All the

parameters for fluorophores were obtained from www.lifetechnologies.com. The calculation for Förster distance was performed using a free online software available at http://pymolwiki.org/index.php/Forster_distance_calculator. Using this Förster distance, the maximum change in the distance between the FRET probes upon Ca^{2+} binding is calculated as 1.3 Å.

Chemical Synthesis

General Methods

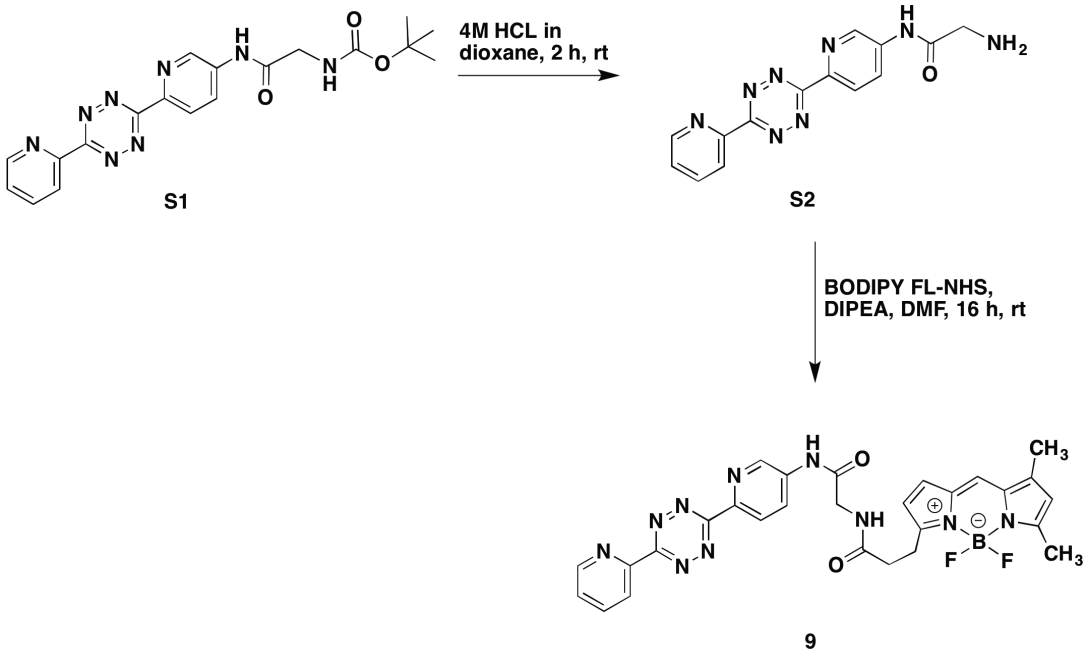
Analytical thin-layer chromatography (TLC) was carried out on silica 60F-254 plates and flash chromatography was carried out on silica gel 60 (230-400 mesh). The spots on the TLC were visualized by UV light and /or ninhydrin and/or Vanillin staining. ESI-MS was acquired using an Agilent 1200 LC-MS system with a 6130 Quadrupole spectrometer. The solvent system used for the LC-MS was 0.2% formic acid in water as buffer A and 0.2% formic acid in acetonitrile (MeCN) as buffer B. For analysing small molecules, Phenomenex Jupiter C18 column (150 x 2 mm, 5 μm) was used for liquid chromatography and mass spectra were acquired in both positive and negative modes. Semi-preparative HPLC purification was carried out using Varian PrepStar/ProStar HPLC system with automated fraction collector. Phenomenex C18 column (250 x 30 mm, 5 μm) column was used to separation and compounds were identified by UV absorbance at 191 nm.

All solvents and chemical reagents were purchased from commercial suppliers and used without further purification. Bicyclo[6.1.0]non-4-yn-9-ylmethanol (BCN-OH, exo/endo mixture ~2/1) was purchased from SynAffix, Netherlands. The ligand used

in Cu-catalyzed click reaction, Tris-(hydroxypropyltriazolylmethyl)amine (THPTA) was synthesized using the protocol described in literature[11].

Unnatural amino acids **1**, **4** and **7** were purchased from Bachem, Switzerland and **3** was purchased from Sigma-Aldrich. Unnatural amino acids **2**, **5**, **6** and **8** were synthesized using a procedure reported earlier[12, 5, 13, 14].

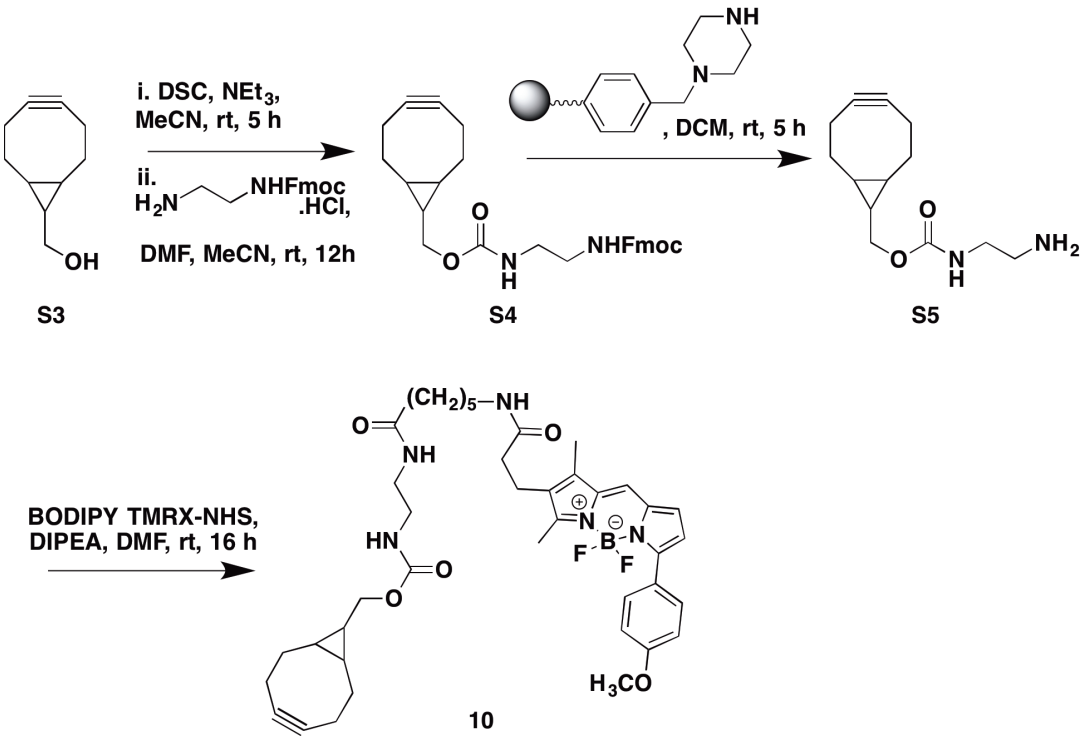
Synthesis of Tetrazine-BODIPY FL (9)



Boc-protected Tetrazine **S1** was synthesized and characterized as previously described.[14] Deprotection was carried out by adding 4M HCl in dioxane (500 µL, 2.0 mmol) to the solution of Tetrazine **S1** (8 mg, 0.02 mmol) in DCM (500 µL) and stirring for 2 h at room temperature. Subsequently the solvent was removed under reduced pressure to yield primary amine hydrochloride **S6** as a pink solid (6mg, 0.02 mmol, 100%). The compound was directly used in the next step without any further purification.

BODIPY FL succinimidyl ester (5mg, 0.013 mmol, Invitrogen) and N,N-Diisopropylethylamine, DIPEA (50 µl, 2.8 mmol) were added to the solution of Tetrazine-amine **S2** (6mg, 0.02 mmol) in dry DMF (1 mL). The reaction mixture was stirred at room temperature for 16 h and subsequently, diluted with 4 mL of water. The product was purified by semi-preparative reverse phase HPLC using a gradient from 10% to 90% of buffer B in buffer A (buffer A: H_2O; buffer B: acetonitrile). The identity and purity of the tetrazine-BODIPY FL conjugate **9** was confirmed by LC-MS. ESI-MS: $[M-H]^-$, calcd. 581.4, found 581.2.

Synthesis of bicyclononyne-BODIPY-TMR-X (10)



To a stirring solution of Bicyclo[6.1.0]non-4-yn-9-ylmethanol (BCN-OH, exo/endo mixture ~2/1) **S3** (600 mg, 4 mmol) in MeCN (10 ml), triethylamine (2 ml, 14.3 mmol) and N,N'-disuccinimidyl carbonate (1.3 g, 5 mmol) were added at 0°C. The reaction mixture was warmed to room temperature and stirred for 5 h. After 5 h, DMF

(100 ml) and mono-Fmoc ethylene diamine hydrochloride (2 g, 6.3 mmol) were added. The reaction was stirred for another 12 h, then diluted with Et_2O (200 ml), and subsequently washed with H_2O (2x100ml) and 2% HCl (2x100ml). The organic layer was dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. The crude was purified by silica gel chromatography (0-5% MeOH in DCM) to yield BCN-NHFmoc **S4** (1.4g, 78% yield over two steps).

Compound **S4** (500 mg, 1.1 mmol) was dissolved in DCM (20ml) and stirred with polymer-bound piperazine (3 g, 3 mmol, 200-400 mesh, 1.0-2.0 mmol/g loading, 2% cross-linked with divinylbenzene) for 4 h at room temperature. Subsequently, the polymer support was removed by filtration and washed with DCM. The filtrate was concentrated under reduced pressure and used directly in the next step assuming 100% yield. The identity and purity of the BCN-NH_2 **S5** was confirmed by LC-MS. ESI-MS: $[M-H]^+$, calcd. 237.3, found 237.3.

BODIPY-TMRX N-hydroxysuccinimide ester (5mg, 0.0082 mmol, Invitrogen) and N,N-Diisopropylethylamine, DIPEA (100 µl, 5.7 mmol) were added to the solution of BCN-NH_2 **S5** (3.5 mg, 0.015 mmol) in dry DMF (1 mL). The reaction mixture was stirred at room temperature for 16 h and diluted with 20 ml of H_2O. The product was purified by semi-preparative reverse phase HPLC using a gradient from 10% to 90% of buffer B in buffer A (buffer A: H_2O; buffer B: acetonitrile). The identity and purity of the bicyclononyne-BODIPY-TMRX conjugate **10** was confirmed by LC-MS. ESI-MS: $[M-H]^-$, calcd. 728.6, found 728.3.

Supplementary Figure Legends

Supplementary Figure S1. Strategy and design of anticodon stem loops for Pyl tRNA(N8)$_{XXXX}$ libraries. The anticodon stem loops shown were strategically chosen

because they have either been used previously for unnatural amino acid incorporation, or because we anticipate- based on the known binding determinants of synthetase - tRNA interactions- that they may be minimally mis-aminoacylated by endogenous synthetases. Following selection on tRNA $(N8)_{UCUA}$ we found U33 and A37 were conserved and therefore changed the library designs for other tRNAs accordingly.

Supplementary Figure S2. The anticodon stem loop sequences of evolved Pyl tRNA(N8)$_{XXXX}$ library members resulting from the selections. a. Four distinct evolved Pyl $tRNA_{UCUA}$ sequences were discovered in the 60 sequenced clones from the selection. **b.** Nine selected Pyl $tRNA_{UCCU}$. 80 clones were picked from selection plate, among which 59 (74%) grew on Chloramphenicol plates without **1** and were not considered further. 6 (8%) didn't grow on Chloramphenicol plates with or without **1** and were not considered further. 15 (19%) grew only on Chloramphenicol plates with **1** but not on Chloramphenicol plates without **1**. These 15 clones were sequenced and revealed 9 different Pyl $tRNA_{UCCU}$ sequences. **c.** Six selected Pyl $tRNA_{UACU}$. 60 clones were picked from selection plate, among which 28 (47%) did not grow on Chloramphenicol plates with or without **1** and were not considered further. 32 out of the 60 (53%) clones grew only on Chloramphenicol plates with **1** but not on Chloramphenicol plates without **1**. These 32 clones were sequenced and revealed 6 different Pyl $tRNA_{UCCU}$ sequences. **d.** Five distinct selected Pyl $tRNA_{CUAG}$ from a total of 59 sequenced clones. The numbers under each Pyl tRNA sequence represent the number of times such a sequence was independently isolated. Clone A in **panels a-d** corresponds to the "evolved Pyl $tRNA_{XXXX}$" in the main text.

Supplementary Figure S3. Evolved Pyl tRNA(N8)$_{XXXX}$ library members enhance the quadruplet decoding of O-*cat*$_{111XXXX}$. The evolved Pyl tRNA$_{UCUA}$ clones (Pyl tRNA$_{UCUA}$ clone A-D, **panel a**), evolved Pyl tRNA$_{UCCU}$ clones (Pyl tRNA$_{UCCU}$ clone A-I, **panel b**), evolved Pyl tRNA$_{UACU}$ clones (Pyl tRNA$_{UACU}$ clone A-F, **panel c**), evolved Pyl tRNA$_{CUAG}$ clones (Pyl tRNA$_{CUAG}$ clone A-E, **panel d**) enhance the incorporation of unnatural amino acids by ribo-Q1 in response to quadruplet codons as measured by survival on increasing concentrations of chloramphenicol (Cm). If the Pyl tRNA$_{XXXX}$ can efficiently decode the quadruplet codon XXXX in O-*cat*$_{111XXXX}$ and the Pyl tRNA$_{XXXX}$ is only a substrate of its cognate PylRS but not any of the twenty endogenous natural aminoacyl tRNA synthetases, full length CAT is only produced in the presence of **1** and not produced in the absence of **1**. This results in survival on chloramphenicol plates only in the presence of **1** but not in the absence of **1**. If the Pyl tRNA$_{XXXX}$ can efficiently decode the quadruplet codon XXXX in O-*cat*$_{111XXXX}$ but the Pyl tRNA$_{XXXX}$ is mis-aminoacylated by any of the twenty endogenous natural aminoacyl-tRNA synthetases, full length CAT is produced in the absence of **1**, corresponding to survival on chloramphenicol plates in the absence of **1**. If the Pyl tRNA$_{XXXX}$ cannot efficiently decode the quadruplet codon XXXX in O-*cat*$_{111XXXX}$, full length CAT is not produced with or without **1**. This results in death on chloramphenicol plates with or without **1**. Clone A in **panels a-d** corresponds to the evolved Pyl tRNA$_{XXXX}$ in the main text. While the evolved CTAG and AGTA decoding tRNAs yielded lower levels of absolute chloramphenicol resistance than other selected tRNAs their activity in recombinant protein expression was comparable to that of other selected tRNAs.

Supplementary Figure S4. Evolved Pyl tRNA(N8)$_{XXXX}$ library members direct efficient unnatural amino acids incorporation in response to quadruplet codons

via orthogonal translation. a. Schematic showing that the incorporation of an unnatural amino acid in response to a specific quadruplet codon XXXX in the linker region of the O-*gst-cam* gene leads to expression of a full length GST-CaM fusion protein. Incorporation of a natural amino acid in response to the first three bases in the quadruplet codon will lead to a frame shift and premature termination, leading to expression of the GST protein only. The ratio between the full-length GST-CaM protein and the truncated GST protein therefore provides a measure of quadruplet decoding efficiency. **b.** Incorporating **4** with *Mj*AzPheRS*/tRNA$_{UCCU}$. **c.** Enhanced incorporation of unnatural amino acids in response to quadruplet codons by evolved Pyl tRNA$_{UCUA}$, Pyl tRNA$_{UCCU}$, Pyl tRNA$_{UACU}$, Pyl tRNA$_{CUAG}$ when compared to the corresponding transplant Pyl tRNA$_{XXXX}$ measured by recombinant protein expression using ribo-Q1, O-*gst-cam$_{1XXXX}$*. The quadruplet codons TAGA, AGGA, AGTA and CTAG replace the first ATG codon of the CaM open reading frame in the O-*gst-cam$_{1XXXX}$* genes. **d**. Repeat of the experiment in **Supplementary Figure 4c**, but using a wild-type ribosome and the indicated mRNA with wild-type ribosome binding site.

The incorporation of **1** with PylRS/evolved tRNA$_{XXXX}$ using Ribo-Q1 and *O-gst-cam$_{1XXXX}$* was more efficient than the incorporation at the same sites using the wild-type ribosome. Moreover, the increase in the efficiency of decoding the evolved Pyl tRNA$_{CUAG}$ and evolved Pyl tRNA$_{AGGA}$ on ribo-Q1 versus the wild-type ribosome was greater than any increase in efficiency of decoding the transplant Pyl tRNA$_{CUAG}$ or transplant Pyl tRNA$_{AGGA}$ on ribo-Q1 versus the wild-type ribosome. This demonstrates that tRNAs selected for decoding on the orthogonal ribosome can be more specific for the orthogonal ribosome than tRNAs that are not specifically evolved to work with the orthogonal ribosome.

e. Diverse unnatural amino acids are efficiently incorporated in response to quadruplet codons using PylRS and the evolved Pyl $tRNA_{XXXX}$ with orthogonal translation. Incorporation of amino acids **2** and **3** in response to quadruplet codon TAGA, AGGA, AGTA, CTAG in O-*gst-cam*$_{1XXXX}$ using evolved Pyl $tRNA_{UCUA}$, Pyl $tRNA_{UCCU}$, Pyl $tRNA_{UACU}$, Pyl $tRNA_{CUAG}$ and ribo-Q1. The incorporation and quadruplet decoding efficiency is significantly improved compared to the system using *Mj*AzPheRS*/$tRNA_{UCCU}$ shown in **panel a**. The molar percentage of each band was calculated based on band intensities on Coomassie Blue (InstantBlue, Expedeon) stained gels and molecular weights of full length and truncated proteins.

Supplementary Figure S5. Single unnatural amino acid incorporation in response to quadruplet codons using different systems. a. Effect of increased PylRS copy number on the efficiency of quadruplet decoding using evolved Pyl $tRNA_{CUAG}$ and ribo-Q1. **b**. Incorporation of **1** in response to stop codon UAA by PylRS/$tRNA_{UUA}$, and to stop codon UGA by PylRS/$tRNA_{UCA}$, and to the AGGA codon by M7 Pyl $tRNA_{UCCU}$, measured by recombinant protein expression using wild-type ribosome *gst-cam*$_{1XXX(X)}$. The codons TAA, TGA, and AGGA replace the first ATG codon of the CaM open reading frame in the *gst-cam*$_{1XXXX}$ genes. Since the efficiency of decoding may be context and expression system dependent, it is possible that these Pyl tRNA variants will be more efficient in other contexts.

Supplementary Figure S6. Systematic improvements in orthogonal translation system for incorporating multiple unnatural amino acids. a. Schematic of incorporating multiple unnatural amino acids via evolved orthogonal translation. **b**. Schematic showing the reporter gene O-*gst-cam*$_{1AGGA+40TAG}$ and the tRNAs (*Mj*tRNA$_{UCCU}$ and PyltRNA$_{CUA}$) that were used in the experiment shown in **Figure 3a** and **Supplementary Figure S7a** for site-specific incorporation of **1** and **4** using ribo-Q1. AGGA replaces the 1st codon and UAG replaces the 40th codon in the *cam* open reading frame of O-*gst-cam* to make O-*gst-cam*$_{1AGGA+40TAG}$. Decoding of both the AGGA and TAG codons by ribo-Q1 produces full length Gst-CaM while failure to decode either of these codons leads to premature termination of the polypeptide. **c**. Schematic showing the reporter gene O-*gst-cam*$_{1AGGA+40TAG}$ and the tRNAs (*Mj*tRNA$_{UCCU}$ and PyltRNA$_{CUA}$) that were used in the experiment shown in **Figure 3b** and **Supplementary Figure S7b** for site-specific incorporation of **1** and **4** using ribo-Q1 and the three plasmid system. The tRNAs and the reporter system are the same as in **Figure 3a** and **Supplementary Figure S7a,** but the number of plasmids are reduced from four to three. **d**. Schematic showing the reporter gene O-*gst-cam*$_{1TAG+40AGGA}$ and the tRNAs (*Mj*tRNA$_{CUA}$ and PyltRNA$_{UCCU}$) that were used in the experiment shown in **Figure 3c** and **Supplementary Figure S7c** for site-specific incorporation of **1** and **4** using ribo-Q1 and the three plasmid system. **e**. Schematic showing the reporter gene O-*gst-cam*$_{1TAG+40AGTA}$ and the tRNAs (*Mj*tRNA$_{CUA}$ and evolved PyltRNA$_{UACU}$) that were used in the experiment shown in **Figure 3d** and **Supplementary Figure S7d** for site-specific incorporation of **1** and **4** using ribo-Q1 and the three plasmid system. **e**. Schematic showing the reporter gene O-*gst-cam*$_{1AGGA+40TAGA}$ and two quadruplet decoding tRNAs (*Mj*tRNA$_{UCCU}$ and evolved PyltRNA$_{UCUA}$) that were used in the experiment shown in **Figure 3e** and

Supplementary Figure S7g for site-specific incorporation of **1** and **4** using ribo-Q1 and the three plasmid system.

Supplementary Figure S7. Efficient incorporation of distinct unnatural amino acids into a single polypeptide. a. Site-specific incorporation of **1** and **4** in cells using ribo-Q1 expressed from a constitutively active promoter and a low copy number number plasmid (pSC101 ribo-Q1) [1], O-*gst-cam*$_{1AGGA}$ from the p22 plasmid[1] and the *Mj*AzPheRS*/tRNA$_{UCCU}$ and PylRS/tRNA$_{CUA}$ pairs[1]. AGGA replaces the 1st codon and UAG replaces the 40th codon in the *cam* open reading frame of O-*gst-cam* to make O-*gst-cam*$_{1AGGA+40TAG}$. Decoding of both the AGGA and TAG codons by ribo-Q1 produces full length Gst-CaM, while failure to decode these codons leads to premature termination of polypeptide synthesis. **b**. The site-specific incorporation efficiency of **1** and **4** is improved by reducing the number of plasmids and increasing PylRS copy number. In these experiments the expression of ribo-Q1 rRNA is induced from an RSF vector. **c**. Incorporation of **1** and **4** in cells expressing Ribo Q1 rRNA from an RSF vector, O-*gst-cam*$_{1TAG+40AGGA}$, *Mj*AzPheRS*/tRNA$_{UCCU}$, and PylRS/tRNA$_{CUA}$. **d**. Site-specific incorporation of **1** and **4** in cells expressing Ribo Q1 rRNA from an RSF vector, O-*gst-cam*$_{1TAG+40AGTA}$, *Mj*AzPheRS/tRNA$_{CUA}$ and the new evolved PylRS/tRNA$_{UACU}$. **e**. Site-specific incorporation of **1** and **4** in cells expressing Ribo Q1 rRNA from an RSF vector, O-*gst-cam*$_{1TAG+40CTAG}$, *Mj*AzPheRS/tRNA$_{CUA}$ and the new evolved PylRS/tRNA$_{CUAG}$. **f**. Site-specific incorporation of **1** and **4** in cells expressing Ribo Q1 rRNA from an RSF vector, O-*gst-cam*$_{1TAG+40AGGA}$, *Mj*AzPheRS/tRNA$_{CUA}$ and the new evolved PylRS/tRNA$_{UCCU}$. **g**. Incorporating two distinct unnatural amino acids using two distinct quadruplet codons in cells expressing Ribo Q1 rRNA from an RSF vector, O-*gst-cam*$_{1AGGA+40TAGA}$, AzPheRS*/tRNA$_{UCCU}$ and the new evolved PylRS/tRNA$_{UCUA}$. **h**. Expression of

protein containing two distinct stop codons, UAG and UAA. Cells were transformed with RSF plasmid expressing wild-type rrnB and *gst-cam*$_{1TAG+40TAA}$, SUP plasmid expressing AzPheRS*/tRNA$_{CUA}$ and CDF plasmid expressing PylRS/tRNA$_{UUA}$. No distinct band for the full length protein was observed. **i**. Expression of protein containing two distinct stop codons, UAG and UGA. Cells were transformed with RSF plasmid expressing wild-type rrnB and *gst-cam*$_{1TAG+40TGA}$, SUP plasmid expressing AzPheRS*/tRNA$_{CUA}$ and CDF plasmid expressing PylRS/tRNA$_{UCA}$. No distinct band for the full length protein was observed. The molar percentage of each band was calculated based on band intensities on Coomassie Blue (InstantBlue, Expedeon) stained gel and molecular weights of full length and truncated proteins.

Supplementary Figure S8. Efficient and specific incorporation of a matrix of pairs of unnatural amino acids demonstrates generality. a. Cells contained O-*gst-cam*$_{1TAG+40AGTA}$ and ribo-Q1 rRNA, expressed from an RSF plasmid, the pSUP *Mj*AzPheRS (or relevant active site variant)/tRNA$_{CUA}$ plasmid and the pCDF-PylRS/tRNA$_{UACU}$ plasmid. All combinations of PylRS substrates (**1**-**3**) and MjTyrRS active site variant substrates (**4-7)** were incorporated in 3×4 matrix. We further confirmed the incorporation of distinct unnatural amino acids by electrospray ionization mass spectrometry. The molar percentage of each band was calculated based on band intensities on Coomassie Blue (InstantBlue, Expedeon) stained gels and the molecular weights of full length and truncated proteins. **b.** MALDI mass spectra of recombinant proteins containing two distinct unnatural amino acids. MALDI mass spectra of each recombinant protein with pairs of unnatural amino acids were obtained using BRUKER TOF-TOF mass spectrometer. Each sample was desalted using 7K MWCO Zeba spin desalting columns (Thermo Scientific). Sinapinic acid was employed as the matrix for these measurements. We observed an

additional peak for protein samples with azide-containing unnatural amino acid that correspond to the reduction of an azide to an amine, as previously reported. The observed molecular weights were in good agreement with the expected molecular weights and within the error range for MALDI for all double incorporations. Note that the mass spectrum for MALDI-MS yields molecular weights as $[M+H]^+$.

Supplementary Figure S9. 5a and 8a do not react and 5a can be labelled with 10. a. 5a and **8a** do not react, even over days. Schematic showing reaction between Nα-Boc-protected 4-(6-methyl-*s*-tetrazin-3-yl)aminophenylalanine (**5a**) and 5-norbornene-2-ol (**8a**). Estimation of second order rate constant for reaction between **5a** and **8a**. The reaction between **5a** and **8a** was performed using 50 mM **8a** and 0.5 mM **5a** to ensure pseudo-first order conditions. Progress of the reaction was followed by measuring loss of **5a** using HPLC. Dotted lines in the plot show expected decrease in **5a** concentration with time at 50 mM **8a** under pseudo-first order conditions for three different second order rate constants. We estimate the second order rate constant for reaction between **5a** and **8a** to be $\sim 10^{-5}$ $M^{-1}s^{-1}$. **b.** Determination of second order rate constant for reaction between **5a** and **10a** by UV spectroscopy. The reactions were performed under pseudo-first order conditions at three different concentrations of bicyclononyne alcohol **10a**, viz. 0.5 mM, 1.5 mM and 2.5 mM. In each of these three reactions, at least 100-fold excess of **10a** compared to **5a** was used to ensure pseudo-first order conditions. Loss of **5a** over the course of the reaction was followed by measuring the absorption at 320 nm. For each of the three concentrations of **10a**, the UV signal at 320 nm was plotted against time and the curve was fit to the following equation: $[\mathbf{5a}]/[\mathbf{5a}]_o = \exp(-k't)$. The pseudo-first order rate constants (k′) are plotted in the figure and fitted to a linear curve. The second order rate constant thus obtained is $k = 0.314 \pm 0.010$ M^{-1} s^{-1}. The reactions were performed in 55/45

MeOH/H_2O mixture at 21°C. All measurements were performed in triplicate and error bars represent the standard deviation.

Supplementary Figure S10. Change of FRET in doubly labelled CaM 5-10$_1$-8-9$_{149}$ in response to increasing urea concentration. a. Production of CaM bearing **5** at position 1 and **8** at position 149. CaM**5**$_1$-**8**$_{149}$ was synthesized and purified from cells bearing pRSF ribo-Q1 O-*gst-cam$_{1TAG+149AGTA}$*, pSUP *Mj*TetPheRS/tRNA$_{CUA}$, pCDF NorKRS$_{x3}$/tRNA$_{UACU}$. CaM **5**$_1$-**8**$_{149}$ can only be produced when both **5** and **8** are supplemented in the growth media. **b.** The purified CaM**5**$_1$-**8**$_{149}$ (**Lane 1**) was labelled with bicyclononyne-BODIPY-TMR-X (**10**) yielding CaM**5-10**$_1$-**8**$_{149}$ (**Lane 2**), and tetrazine-BODIPY-FL (**9**) yielding CaM**5**$_1$-**8-9**$_{149}$ (**Lane 3**). The double labelled product CaM**5-10**$_1$-**8-9**$_{149}$ was obtained by sequential labelling with **9** and then **10** (**Lane 4**). Each labelling was performed with 100 equivalents (200 μM) of the dye conjugate and was quantitative as confirmed by electrospray ionization mass spectrometry. Labelling was visualized by fluorescence imaging and led to a mobility shift on SDS-PAGE. **c**. Fluorescence spectra of double labelled CaM**5-10**$_1$-**8-9**$_{149}$ (orange trace), single labelled CaM**5**$_1$-**8-9**$_{149}$ (green trace, donor fluorophore only), and the single labelled CaM**5-10**$_1$-**8**$_{149}$ (red trace, acceptor fluorophore only) were measured following excitation at 485 nm. **d.** Fluorescence spectra of doubly labelled CaM**5-10**$_1$-**8-9**$_{149}$ following donor excitation at wavelength 485 nm in the presence of increasing concentrations of the denaturant urea are measured, normalised and plotted. Fluorescence spectra of single labelled CaM**5**$_1$-**8-9**$_{149}$ (green trace, donor only) is shown as a reference.

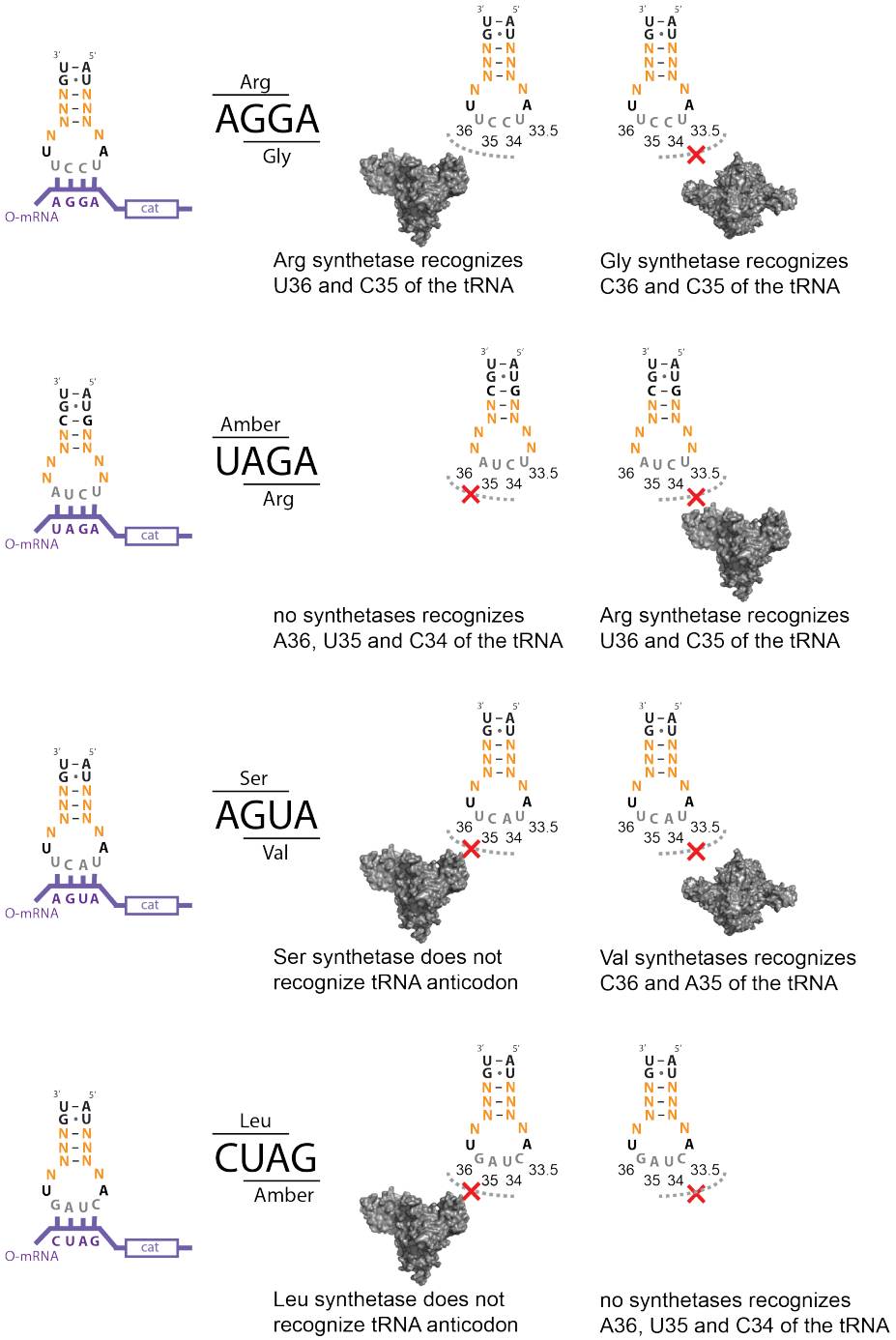
Supplementary Figure S11. Change in FRET of CaM 5-10$_1$-8-9$_{40}$ as a function of Ca^{2+} concentration. a. Production of CaM bearing **5** at position 1 and **8** at position

40. CaM**5**$_1$-**8**$_{40}$ was synthesized and purified from cells bearing pRSF ribo-Q1 O-*gst-cam*$_{1TAG+40AGTA}$, pSUP *Mj*TetPheRS/tRNA$_{CUA}$, pCDF NorKRS$_{.3}$/tRNA$_{UACU}$. CaM **5**$_1$-**8**$_{40}$ can only be produced when both **5** and **8** are supplemented in the growth media. **b.** CaM**5**$_1$-**8**$_{40}$ was synthesized and purified from cells bearing O-*gst-cam*$_{1TAG+40AGTA}$, ribo-Q1, NorKRS$_{.3}$/evolved tRNA$_{UACU}$, AzPheRS/tRNA$_{CUA}$ (**Lane 1**). The purified protein was labelled with tetrazine-BODIPY-FL (**9**) yielding CaM**5**$_1$-**8-9**$_{40}$, in which the norbornene in CaM**5**$_1$-**8**$_{40}$ is labelled with **9** in an inverse electron-demand Diels-Alder reaction (**Lane 3**). Labelling was visualised by fluorescence imaging and led to a mobility shift in SDS PAGE. CaM**5**$_1$-**8-9**$_{40}$ was labelled with 100 equivalents of bicyclononyne-BODIPY-TMR-X (**10**), yielding CaM**5-10**$_1$-**8-9**$_{40}$, in which the encoded tetrazine in CaM**5**$_1$-**8-9**$_{40}$ was labelled via an inverse electron-demand Diels-Alder reaction (**Lane 4**). Labelling was visualized by fluorescence imaging and led to a further mobility shift in SDS-PAGE. Labelling CaM**5**$_1$-**8**$_{40}$ directly with **10** also led to a mobility shift and yielded CaM**5-10**$_1$-**8**$_{40}$ (**Lane 2**). All labelling reactions were quantitative, as confirmed by electrospray ionization mass spectrometry, right hand panel. **c.** CaM**5-10**$_1$-**8-9**$_{40}$ exhibits FRET. Fluorescence spectra of CaM**5-10**$_1$-**8-9**$_{40}$ (orange trace), single labelled CaM**5**$_1$-**8-9**$_{40}$ (green trace, donor fluorophore only), and the single labelled CaM**5-10**$_1$-**8**$_{40}$ (red trace, acceptor fluorophore only) are measured following excitation at 485 nm. **d**. Fluorescence spectra of CaM**5-10**$_1$-**8-9**$_{40}$ as a function of [Ca^{2+}] following excitation at 485 nm. The donor fluorescence increases while the acceptor fluorescence decreases as a function of Ca^{2+}. **e.** Fluorescence spectra of CaM**8-9**$_{40}$ at [Ca^{2+}] = 0 and 30 mM following excitation at 485 nm. This control for the effect of Ca^{2+} concentration on the fluorescence spectra of the donor emission was measured. Unlike the fluorescence spectra of the doubly labelled protein (CaM**5-10**$_1$-**8-9**$_{40}$), a slight decrease in donor fluorescence at [Ca^{2+}]= 30 mM

for singly labelled protein, CaM**8-9**$_{40}$ was observed. The change in FRET on increasing Ca^{2+} concentration are therefore likely to be underestimated. **f**. The calcium dependent FRET is a poor fit to a single component model. The relative donor fluorescence intensity (I') is calculated as $I'=(I_i - I_0)/(I_{MAX} - I_0)$ where I_i is the donor fluorescence intensity at emission wavelength 515 nm from doubly labelled CaM**5-10**$_1$-**8-9**$_{40}$ at a given Ca^{2+} concentration, I_0 is the donor fluorescence intensity at emission wavelength 515 nm at the minimum Ca^{2+} concentration, and I_{MAX} is the donor fluorescence intensity at emission wavelength 515 nm at the maximum Ca^{2+} concentration. The data fits poorly to a single component fitting, $I'= (a_1[Ca^{2+}]/([Ca^{2+}] + K_{D1}))$ with a $R^2 = 0.6534$.

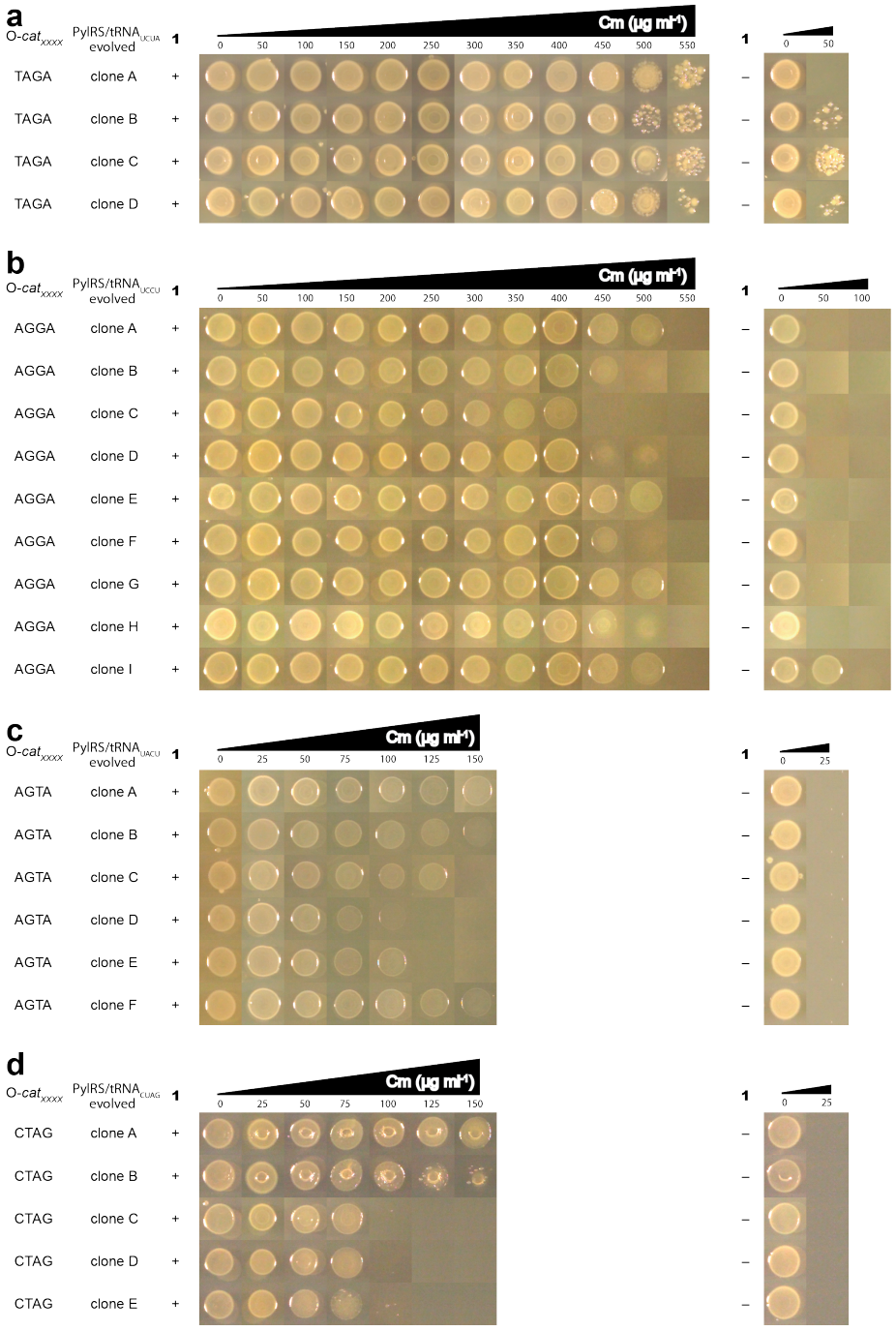
Supplementary Figures

Supplementary Fig. S1

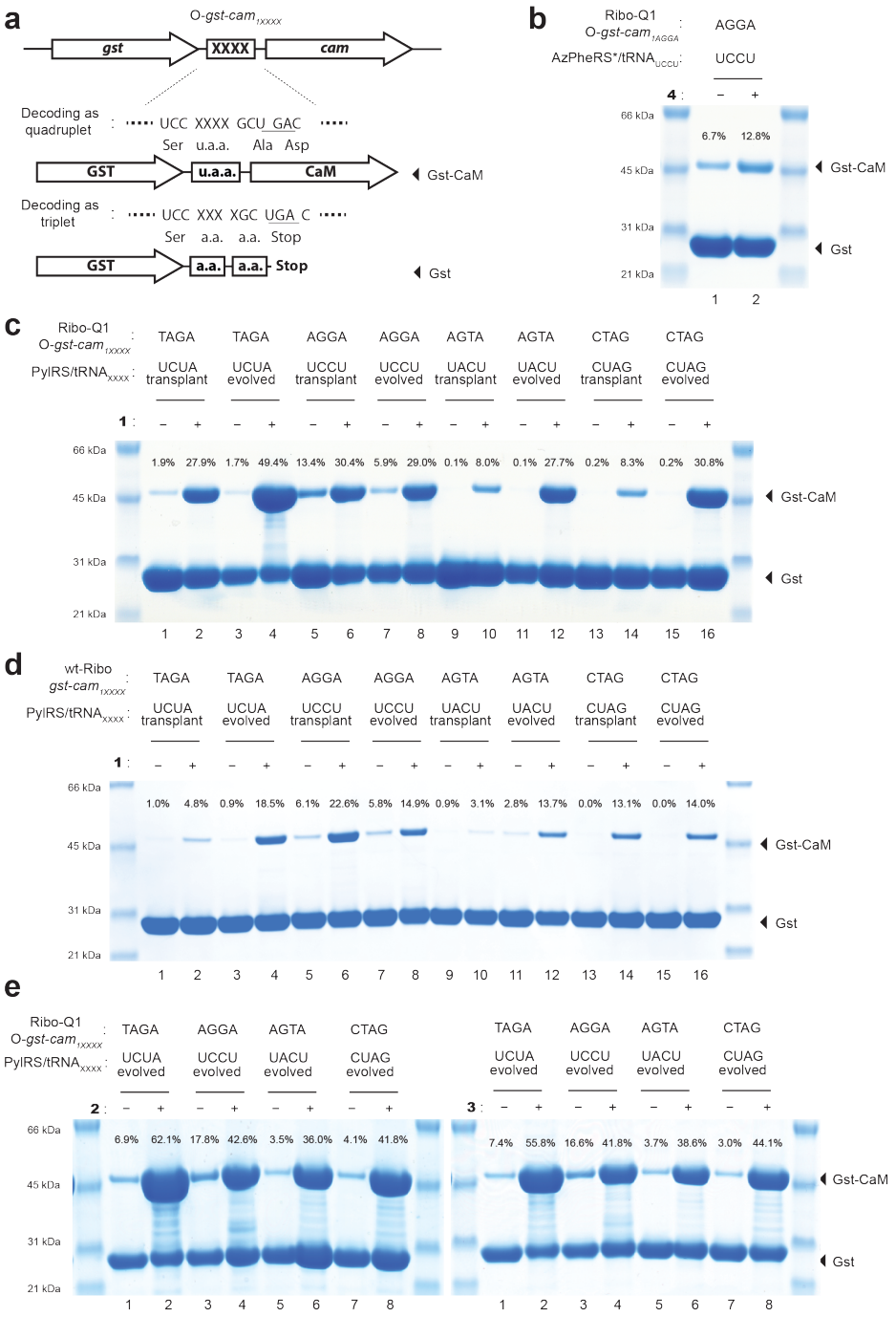


Supplementary Fig. S2

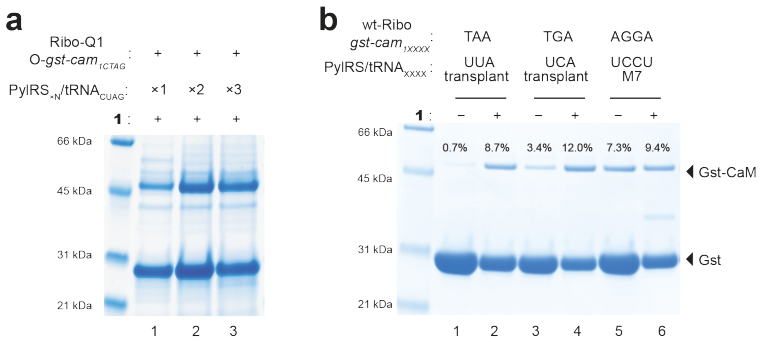
Supplementary Fig. S3



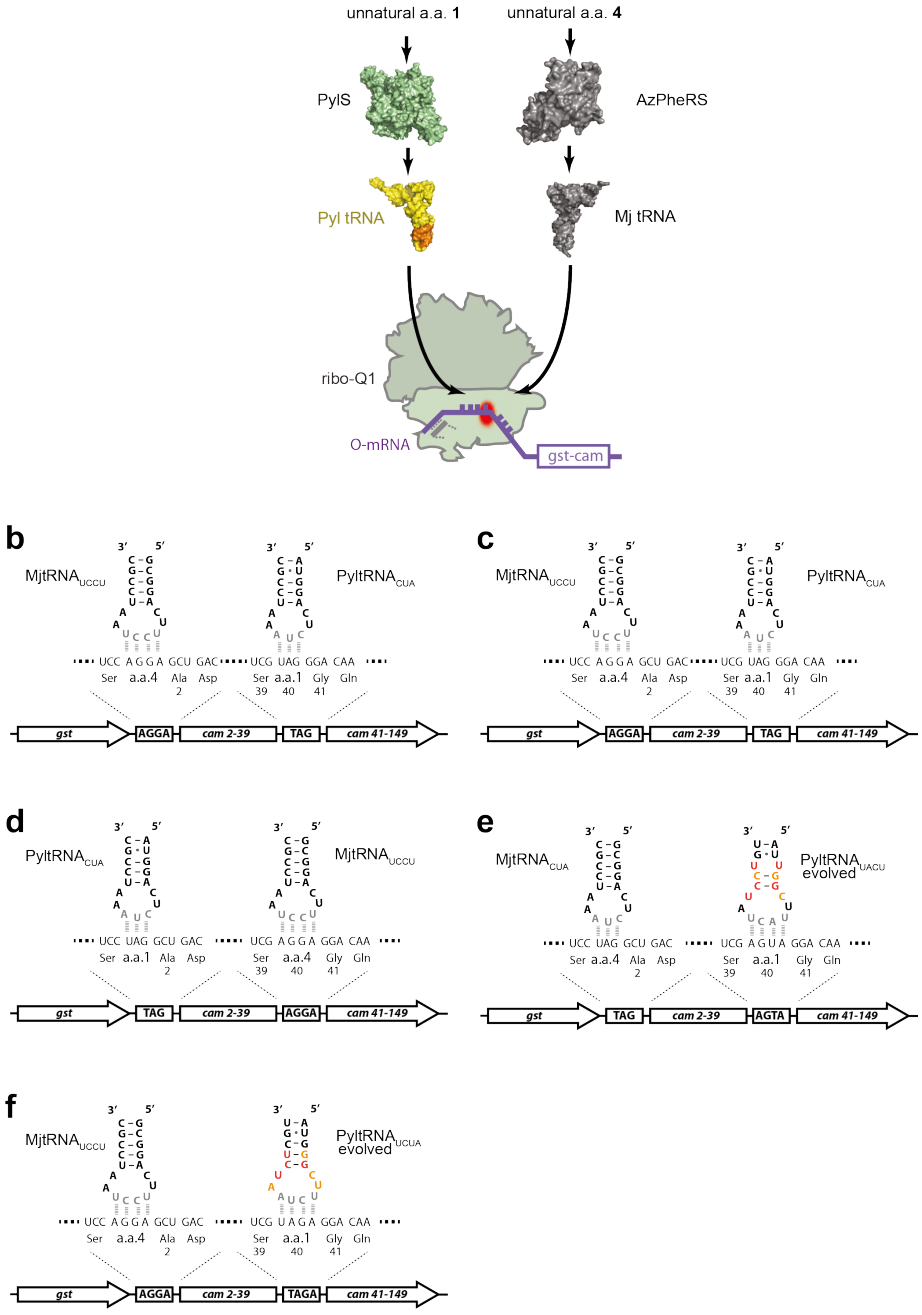
Supplementary Fig. S4



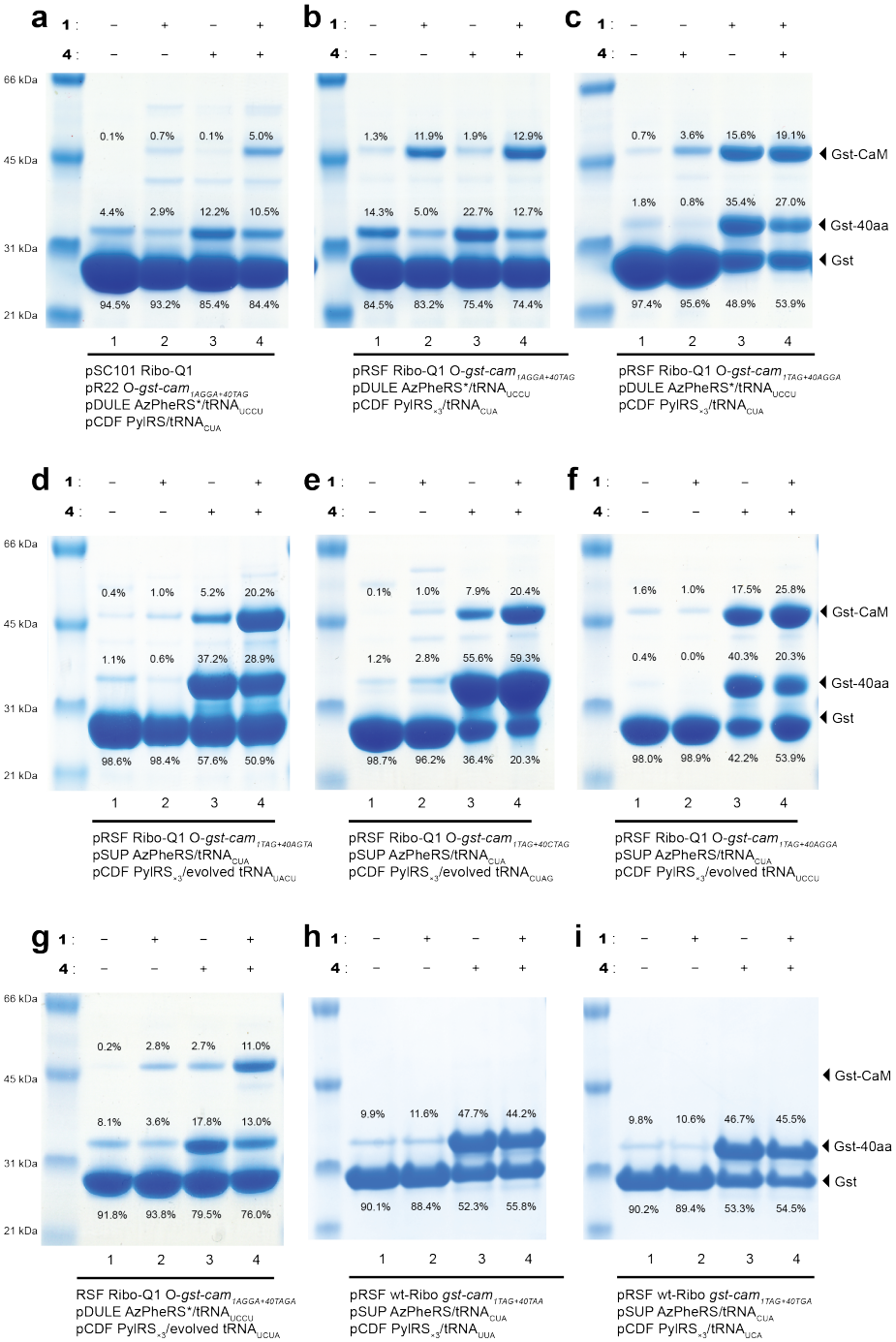
Supplementary Fig. S5



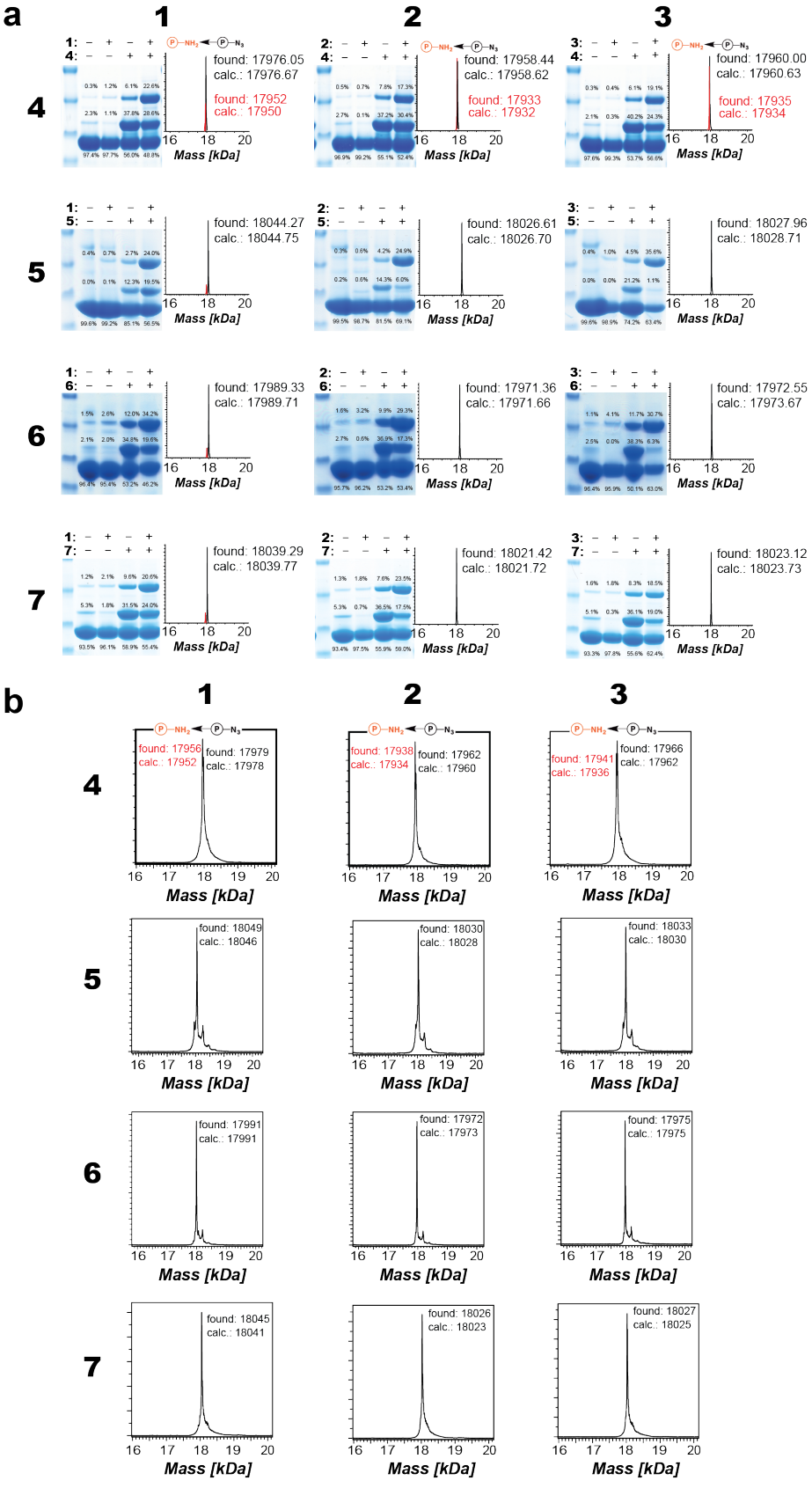
Supplementary Fig. S6



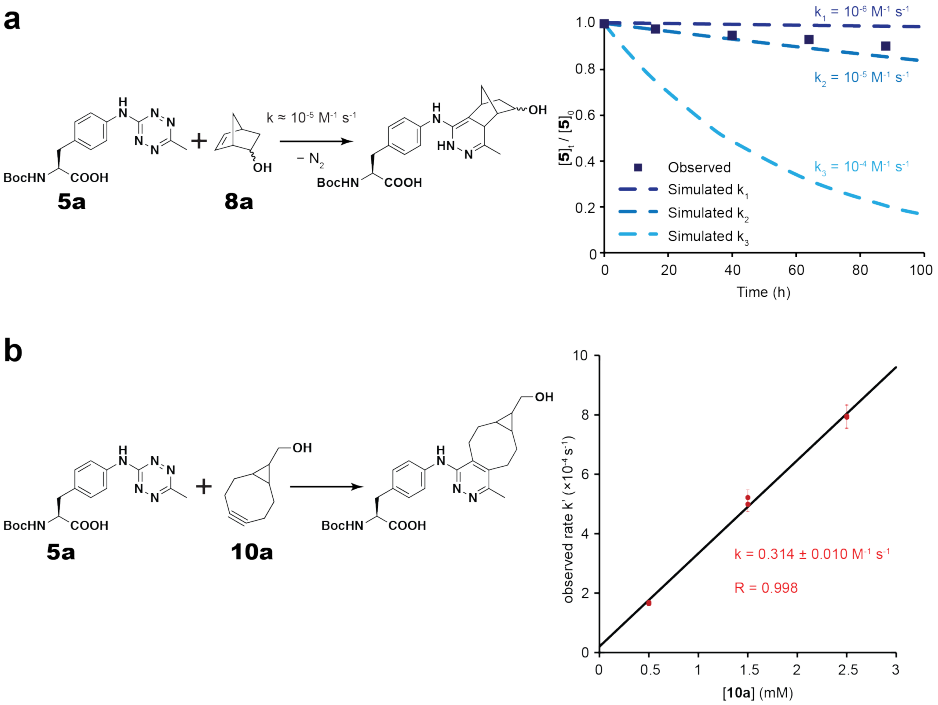
Supplementary Fig. S7



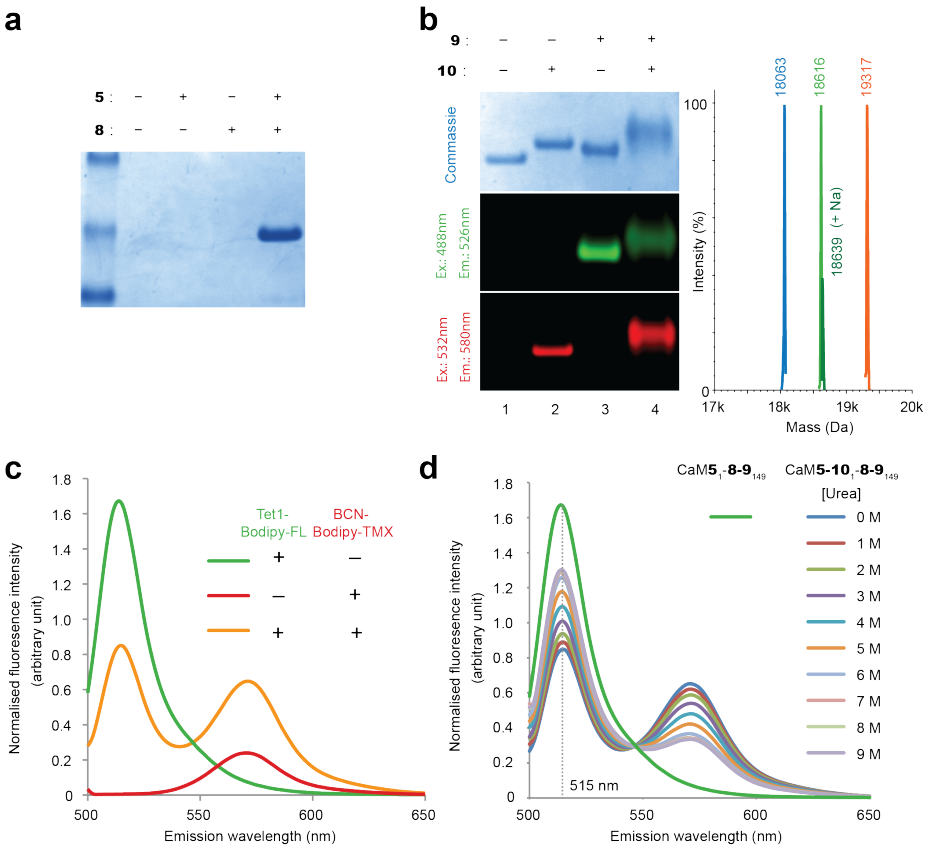
Supplementary Fig. S8



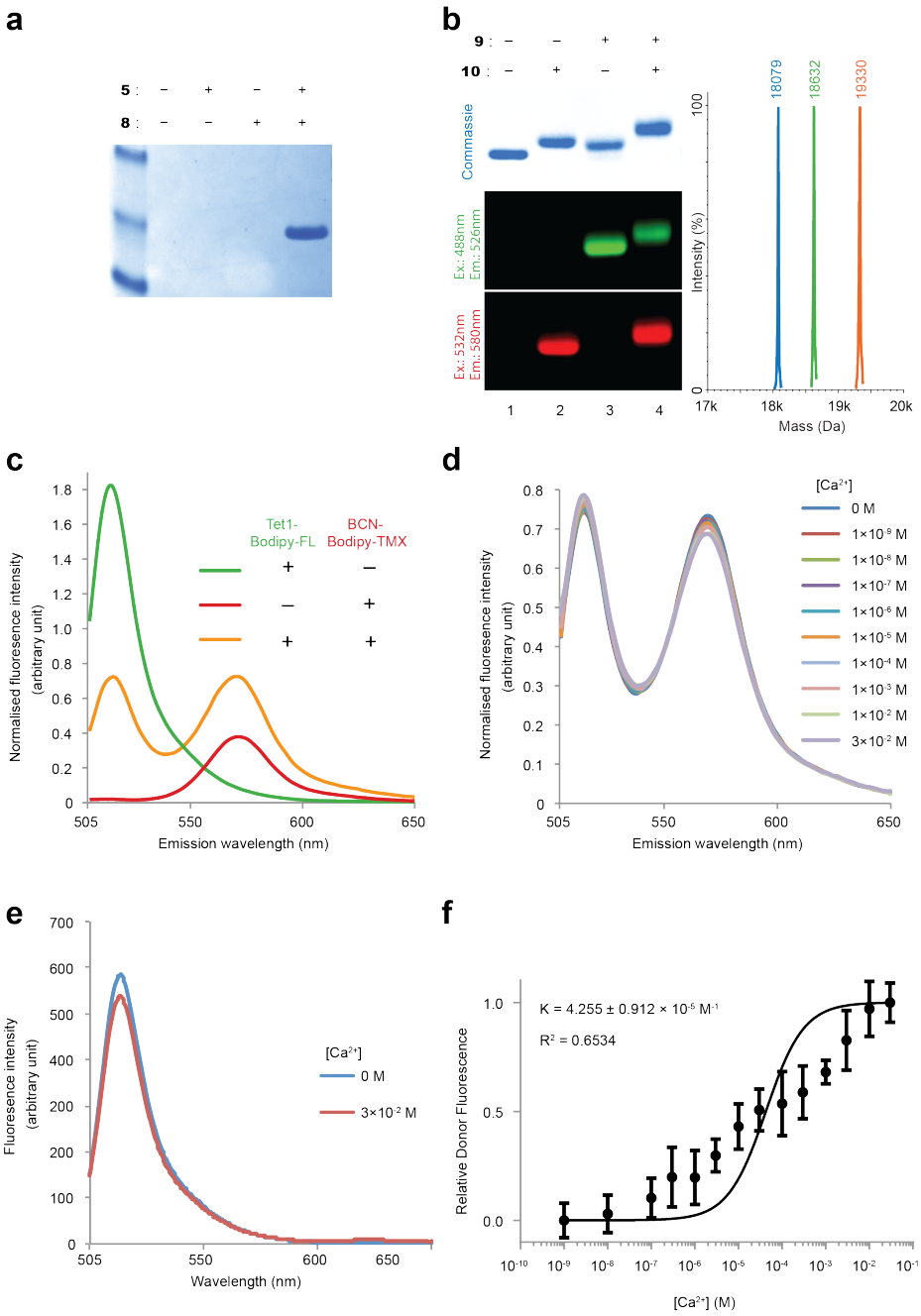
Supplementary Fig. S9



Supplementary Fig. S10



Supplementary Fig. S11



Supplementary Table 1

Primer Name	Primer Sequence (5' to 3')
PylTtUAGAf	CAGAAggtctcaGCTGAACGGATTtagaAGTCCATTCGATCTACA TGATCAGGTTtCCagatctAGC
PylTtUAGAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTtAGGAf	CAGAAggtctcaGCTGAACGGATTaggaAGTCCATTCGATCTACA TGATCAGGTTtCCagatctAGC
PylTtAGGAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTtAGTAf	CAGAAggtctcaGCTGAACGGATTagtaAGTCCATTCGATCTACA TGATCAGGTTtCCagatctAGC
PylTtAGTAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTtCUAGf	CAGAAggtctcaGCTGAACGGATTctagAGTCCATTCGATCTACA TGATCAGGTTtCCagatctAGC
PylTtCUAGr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
MmPylS_Y306Gf	CGATGCTGGCACCGAATCTGggtAATTATCTGCGCAAACTGG A
MmPylS_Y306Gr	TCCAGTTTGCGCAGATAATTaccCAGATTCGGTGCCAGCATC G
MmPylS_Y384Ff	TGGGTGATAGCTGCATGGTGtttGGTGATACCCTGGATGTTAT
MmPylS_Y384Fr	ATAACATCCAGGGTATCACCaaaCACCATGCAGCTATCACCC A
MmPylS_I405Rf	GCAGCGCAGTTGTTGGTCCGcgtCCGCTGGATCGTGAATGG GG
MmPylS_I405Rr	CCCCATTCACGATCCAGCGGacgCGGACCAACAACTGCGCT GC
MmPylS_D318dBamHIf	CACTGCCGGAcCCGATTAA
MmPylS_D318dBamHIr	ATCGGgTCCGGCAGTGC
pCDF1BamHIf	agccaggatccTCGGGagttgtcag
pCDF1743BglIIr	GAAagatctCACCATACCCACGCCGAAACAAG
CaM1UAGf	cgtggatccTAGgctgaccaactgacagaagagc
CaM1UAGr	gttggtcagcCTAggatccacgcggaaccagatcc
CaM1UAGAf	cgtggatccTAGAgctgaccaactgacagaagagc
CaM1UAGAr	gttggtcagcTCTAggatccacgcggaaccagatcc
CaM1AGGAf	cgtggatccAGGAgctgaccaactgacagaagagc
CaM1AGGAr	gttggtcagcTCCTggatccacgcggaaccagatcc
CaM1AGUAf	cgtggatccAGTAgctgaccaactgacagaagagc
CaM1AGUAr	gttggtcagcTACTggatccacgcggaaccagatcc
CaM1CUAGf	cgtggatccCTAGgctgaccaactgacagaagagc
CaM1CUAGr	gttggtcagcCTAGggatccacgcggaaccagatcc
CaM40UAGf	atgaggtcgTAGggacaaaacccaacggaagcag
CaM40UAGr	gttttgtccCTAcgacctcataacggtgccaagttc
CaM40UAGAf	atgaggtcgTAGAggacaaaacccaacggaagcag
CaM40UAGAr	gttttgtccTCTAcgacctcataacggtgccaagttc
CaM40AGGAf	atgaggtcgAGGAggacaaaacccaacggaagcag

CaM40AGGAr	gttttgtccTCCTcgacctcataacggtgccaagttc
CaM40AGUAf	atgaggtcgAGTAggacaaaacccaacggaagcag
CaM40AGUAr	gttttgtccTACTcgacctcataacggtgccaagttc
CaM40CUAGf	atgaggtcgCTAGggacaaaacccaacggaagcag
CaM40CUAGr	gttttgtccCTAGcgacctcataacggtgccaagttc
CaM149UAGf	tgatgacagcaTAGcatcaccatcaccatcactaagcttaattagctg
CaM149UAGr	gtgatggtgatgCTAtgctgtcatcatttgtacaaactcttcgtagt
CaM149UAGAf	tgatgacagcaTAGAcatcaccatcaccatcactaagcttaattagctg
CaM149UAGAr	gtgatggtgatgTCTAtgctgtcatcatttgtacaaactcttcgtagt
CaM149AGGAf	tgatgacagcaAGGAcatcaccatcaccatcactaagcttaattagctg
CaM149AGGAr	gtgatggtgatgTCCTtgctgtcatcatttgtacaaactcttcgtagt
CaM149AGUAf	tgatgacagcaAGTAcatcaccatcaccatcactaagcttaattagctg
CaM149AGUAr	gtgatggtgatgTACTtgctgtcatcatttgtacaaactcttcgtagt
CaM149CUAGf	tgatgacagcaCTAGcatcaccatcaccatcactaagcttaattagctg
CaM149CUAGr	gtgatggtgatgCTAGtgctgtcatcatttgtacaaactcttcgtagt
GstCaMgeneF_KpnI	caGGTACCagcgcccaatacgcaaaccgcctctcc
GstCaMgeneR_KpnI	caGGTACCcaccgcatatggtgcactctcagtacaatc
PylTN8UAGAf	CAGAAggtctcaGCTGAACGNNNNtagaNNNNCATTCGATCTAC ATGATCAGGTTCCCagatctAGC
PylTN8UAGAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTN8AGGAf	CAGAAggtctcaGCTGAACNNNNTaggaANNNNATTCGATCTAC ATGATCAGGTTCCCagatctAGC
PylTN8AGGAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTN8AGUAf	CAGAAggtctcaGCTGAACNNNNTagtaANNNNATTCGATCTACA TGATCAGGTTCCCagatctAGC
PylTN8AGUAr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
PylTN8CUAGf	CAGAAggtctcaGCTGAACNNNNTctagANNNNATTCGATCTACA TGATCAGGTTCCCagatctAGC
PylTN8CUAGr	CAGAAggtctctCAGCCGGGTTAGATTCCCGGGGTTTCCGCCA actag
catD111UAGAf	CGAATACCACtagaGACTTCCGCCAGTTCCTGCAC
catD111UAGAr	TGGCGGAAGTCctaGTGGTATTCGCTCCACAGGCTG
catD111AGGAf	CGAATACCACaggaGACTTCCGCCAGTTCCTGCAC
catD111AGGAr	TGGCGGAAGTCtcctGTGGTATTCGCTCCACAGGCTG
catD111AGTAf	CGAATACCACagtaGACTTCCGCCAGTTCCTGCAC
catD111AGTAr	TGGCGGAAGTCtactGTGGTATTCGCTCCACAGGCTG
catD111CUAGf	CGAATACCACctagGACTTCCGCCAGTTCCTGCAC
catD111CUAGr	TGGCGGAAGTCctagGTGGTATTCGCTCCACAGGCTG
G9_BarnaseF_NotI	TGgcggccgctttcATATCCCTccgcaaATGGCAtagGTTATCAACAC GTTTGACG
BarnaseR_PmeI	acGTTTAAACttaTCTGATTTTTGTAAAGGTCTGATAATGGTCC GTTG
G9_BarUAGAx2F	aagcggccgctttcATATCCCTccgcaaATGGCAtagaGTTATCAACAC GTTTGACGGGGTTGCGGATTATCTTCAG

G9_BarUAGAx2R	TGCCCGGGAGTTTGCCTTCCCTGTTTGAGAAGATGTCTCCG CCGATGCTTTTCCCCGGAGCGACtctaTGCAAGGTTCCCTTTT GATGCCACCCAG
G9_BarAGGAx2F	aagcggccgctttcATATCCCTccgcaaATGGCAaggaGTTATCAACA CGTTTGACGGGGTTGCGGATTATCTTCAG
G9_BarAGGAx2R	TGCCCGGGAGTTTGCCTTCCCTGTTTGAGAAGATGTCTCCG CCGATGCTTTTCCCCGGAGCGACtcctTGCAAGGTTCCCTTTT GATGCCACCCAG
G9_BarAGUAx2F	aagcggccgctttcATATCCCTccgcaaATGGCAagtaGTTATCAACAC GTTTGACGGGGTTGCGGATTATCTTCAG
G9_BarAGUAx2R	TGCCCGGGAGTTTGCCTTCCCTGTTTGAGAAGATGTCTCCG CCGATGCTTTTCCCCGGAGCGACtactTGCAAGGTTCCCTTTT GATGCCACCCAG
G9_BarCUAGx2F	aagcggccgctttcATATCCCTccgcaaATGGCActagGTTATCAACAC GTTTGACGGGGTTGCGGATTATCTTCAG
G9_BarCUAGx2R	TGCCCGGGAGTTTGCCTTCCCTGTTTGAGAAGATGTCTCCG CCGATGCTTTTCCCCGGAGCGACctagTGCAAGGTTCCCTTT TGATGCCACCCAG

References

1. Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J.W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**, 441-444 (2010).

2. Kaya, E. et al. A genetically encoded norbornene amino acid for the mild and selective modification of proteins in a copper-free click reaction. *Angew. Chem. Int. Ed.* **51**, 4466-4469 (2012).

3. Taylor, D.A., Sack, J.S., Maune, J.F., Beckingham, K. & Quiocho, F.A. Structure of a recombinant calmodulin from *Drosophila melanogaster* refined at 2.2-A resolution. *J. Biol. Chem.* **266**, 21375-21380 (1991).

4. Ryu, Y. & Schultz, P.G. Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. *Nat. Methods* **3**, 263-265 (2006).

5. Seitchik, J.L. et al. Genetically encoded tetrazine amino acid directs rapid site-specific *in vivo* bioorthogonal ligation with trans-cyclooctenes. *J. Am. Chem. Soc.* **134**, 2898-2901 (2012).

6. Deiters, A. & Schultz, P.G. *In vivo* incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg. Med. Chem. Lett.* **15**, 1521-1524 (2005).

7. Chin, J.W., Martin, A.B., King, D.S., Wang, L. & Schultz, P.G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A* **99**, 11020-11024 (2002).

8. Neumann, H., Peak-Chew, S.Y. & Chin, J.W. Genetically encoding Nε-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **4**, 232-234 (2008).

9. Kajihara, D. et al. FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids. *Nat. Methods* **3**, 923-929 (2006).

10. Crouch, T.H. & Klee, C.B. Positive cooperative binding of calcium to bovine brain calmodulin. *Biochemistry* **19**, 3692-3698 (1980).

11. Hong, V., Presolski, S.I., Ma, C. & Finn, M.G. Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed.* **48**, 9879-9883 (2009).

12. Nguyen, D.P. et al. Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins viaa pyrrolysyl-tRNA synthetase/$tRNA_{CUA}$ pair and click chemistry. *J. Am. Chem. Soc.* **131**, 8720-8721 (2009).

13. Deiters, A. et al. Adding amino aids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **125**, 11782-11783 (2003).

14. Lang, K. et al. Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat. Chem.* **4**, 298-304 (2012).