Supplementary Material

Proteome labelling and protein identification in specific tissues and at specific developmental stages in an animal

Thomas S. Elliott[1,2], Fiona M. Townsley[1,2], Ambra Bianco[1,2], Russell J. Ernst[1,3], Amit Sachdeva[1,3], Simon J. Elsässer[1,3], Lloyd Davis[1,3], Kathrin Lang[1,3], Rudolf Pisa[1,4], Sebastian. Greiss[1], Kathryn S. Lilley[5] & Jason W. Chin[1,4]

[1]Medical Research Council Laboratory of Molecular Biology, Cambridge, England, UK.

[4]Department of Chemistry, University of Cambridge, Cambridge, England, UK

[5]Department of Biochemistry, University of Cambridge, Cambridge, England, UK

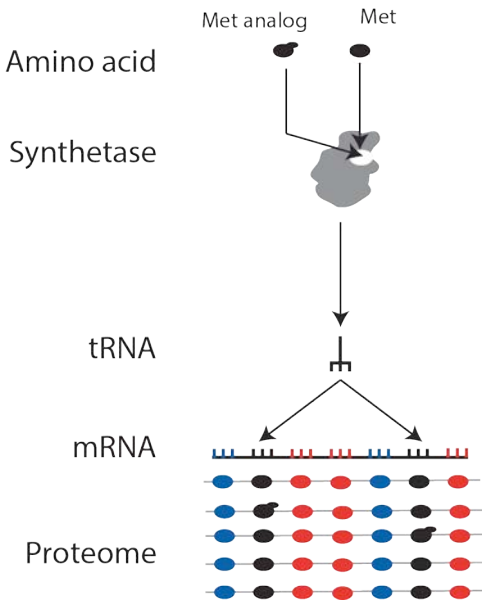
[2]These authors contributed equally.

[3]These authors contributed equally.

E-mail: chin@mrc-lmb.cam.ac.uk

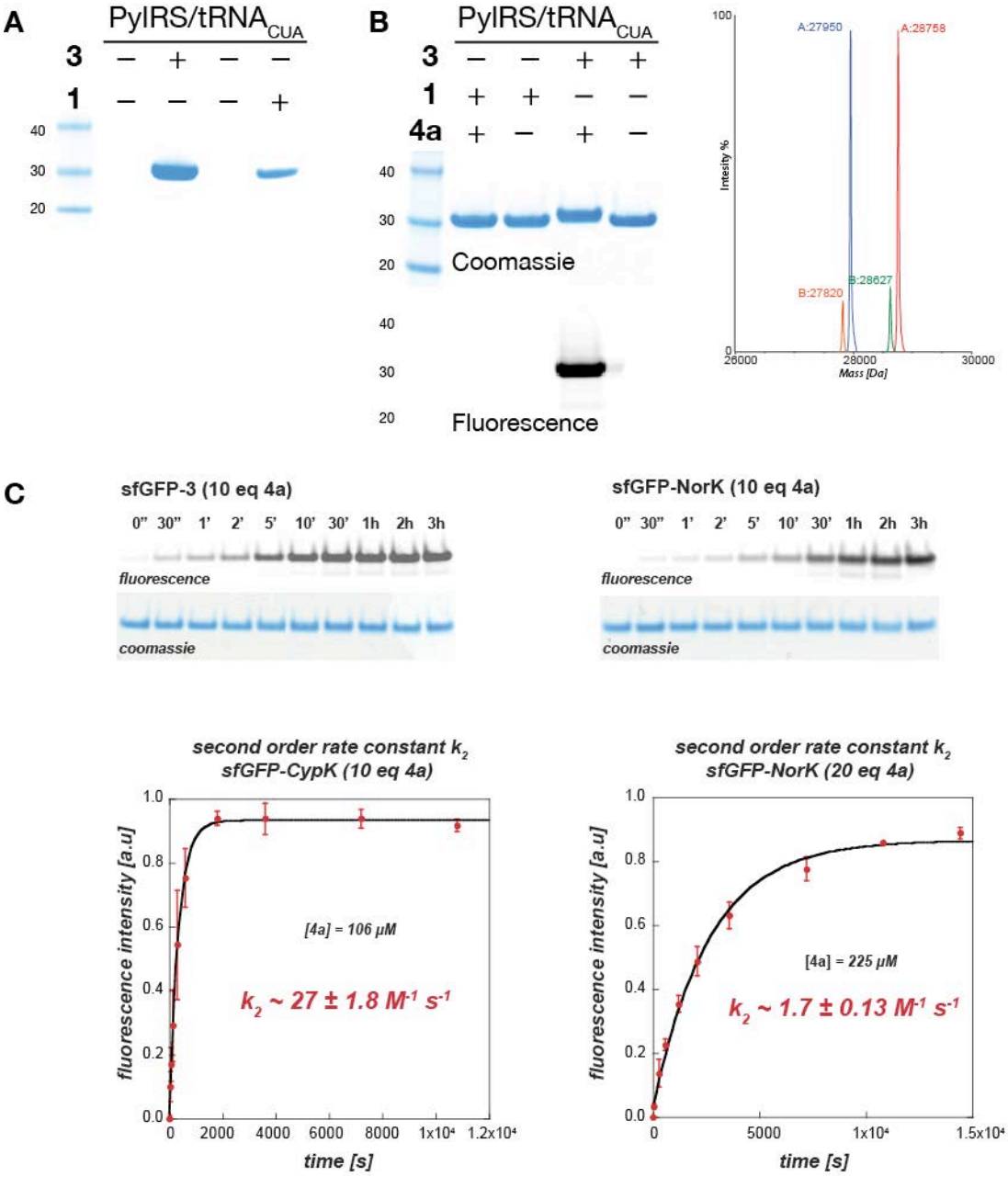
Supplementary Figures

Supplementary Figure 1. Current methods for proteome labelling with unnatural amino acids: selective pressure incorporation



A close analog of a Methionine (natural amino acids are depicted as ovals) competes at the methionyl-tRNA synthetase active site and is statistically incorporated in response to Met codons in the proteome. The approach is best performed under starvation (minimal media) to minimize natural amino acid competition at the active site and is limited to Met codons and close structural analogs of natural amino acids. Variants of the approach have been reported, some using point mutants of the natural synthetase, that expand the chemical scope.

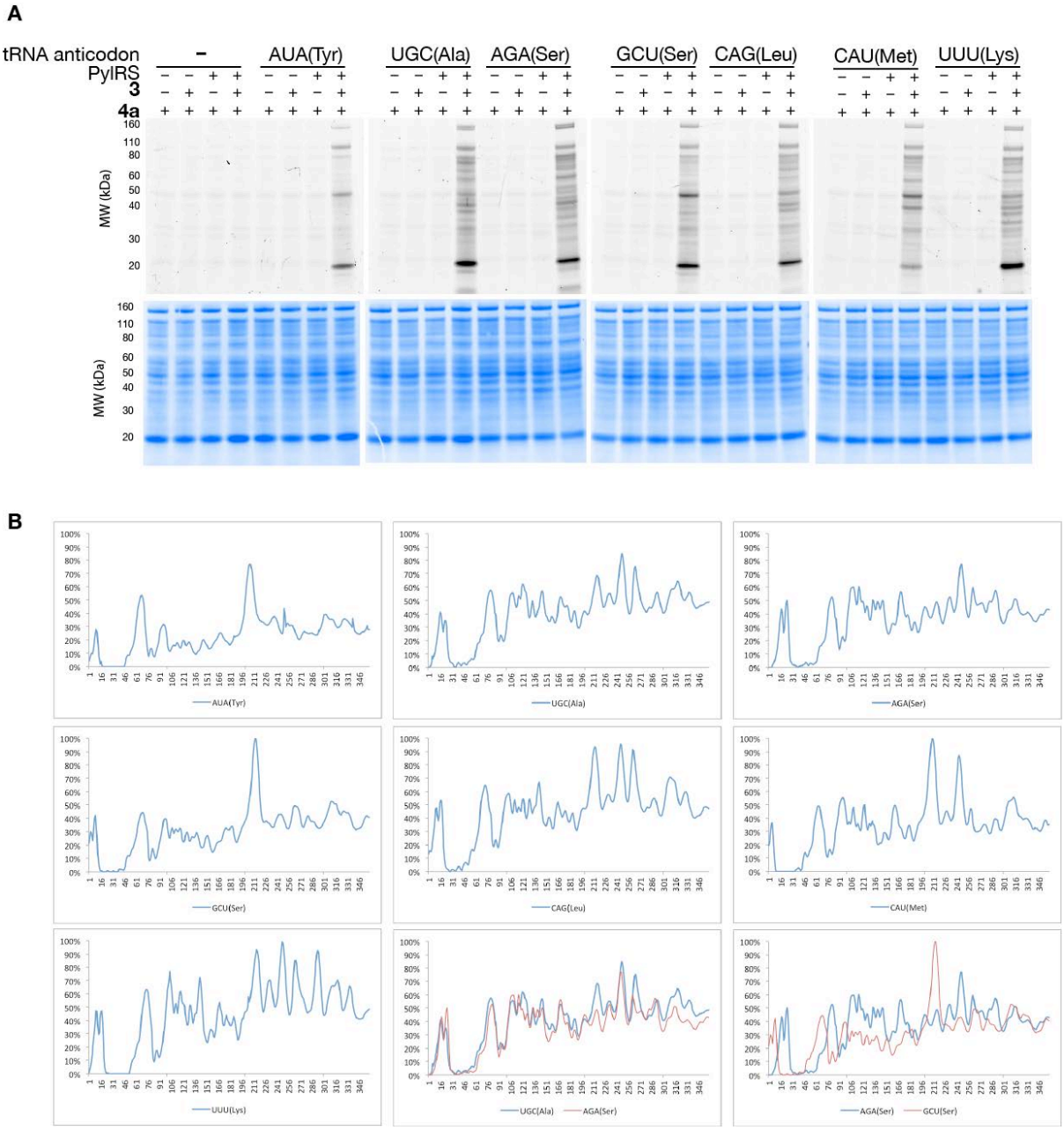
Supplementary Figure 2. Quantitative site-specific incorporation of 3 into proteins expressed in *E. coli* and its rapid and quantitative labelling with tetrazine probes



A. The PylRS/tRNA$_{CUA}$ pair directs efficient, site-specific incorporation of **3** into sfGFP bearing an amber stop codon at position 150. Incorporation of **3** is more efficient than **1** a well-established excellent substrate for the PylRS/tRNA$_{CUA}$ pair. **B.** Specific and quantitative labelling of 2 nmol sfGFP bearing **3** with 10 equivalents of tetrazine fluorophore **4a**. ESI-MS

analysis of sfGFP-**3** purified from *E. coli* grown with 1 mM **3** bearing the PylRS/$tRNA_{CUA}$ pair and SfGFP150TAG confirms the incorporation of **3**. sfGFP150-**3**: Expected mass: 27951.5 Da, Found mass: 27950 ± 1.0 Da, minor peak 27820 corresponding to loss of N-terminal methionine. Labelling sfGFP150-**3** with **4a** is quantitative, as judged by ESI-MS of the labelling reaction. Expected mass: 28758.4 Da, Found mass: 28758 ± 1.0 Da, minor peak 28627 corresponds to loss of N-terminal methionine. **C.** Determining the rate constant for labelling of sfGFP-**3** (10.6 μM, sfGFP incorporating **3** at position 150), with 10 equivalents of **4a.** 2 nmol of purified sfGFP-**3**, (10.6 μM in 20 mM Tris-HCl, 100 mM NaCl, 2 mM EDTA, pH 7.4) were incubated with 20 nmol of tetrazine-dye conjugate **4a** (10 μl of a 2 mM solution in DMSO). At different time points 8 μL aliquots were taken from the solution and quenched with a 700-fold excess of BCN and plunged into liquid nitrogen. Samples were mixed with NuPAGE LDS sample buffer supplemented with 5 % β-mercaptoethanol, heated for 10 min to 90ºC and analyzed by 4-12% SDS page. The amounts of labelled proteins were quantified by scanning the fluorescent bands with a Typhoon Trio phosphoimager (GE Life Sciences). Bands were quantified with the ImageQuant™ TL software (GE Life Sciences) using rubber band background subtraction. The rate constant was determined by fitting the data to a single-exponential equation. The calculated observed rate k' was divided by the concentration of **4a** to obtain rate constant k for the reaction. Measurements were done in triplicate. All data processing was performed using Kaleidagraph software (Synergy Software, Reading, UK). For comparison the rate of labelling sfGFP bearing Nε-5-norbornene-2-yloxycarbonyl-L-lysine (NorK), a known substrate for PylRS, was determined in a similar way using 11.25 μM sfGFP bearing NorK at position 150 (SfGFP-NorK) and 20 equivalents of **4a**.

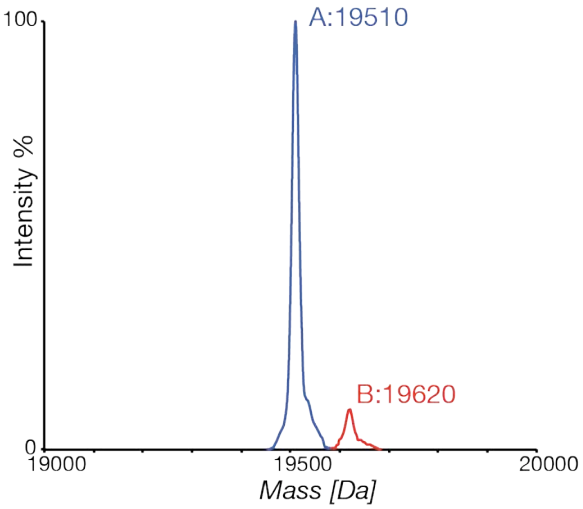
Supplementary Figure 3. SORT-M enables codon specific proteome tagging and labelling in *E. coli*



A. Proteome labelling with **3** via the indicated PylRS/$tRNA_{XXX}$ pair**.** Cells contained two plasmids, one encoding MbPylRS, the other encoding T4 lysozyme and the indicated $tRNA_{XXX}$. Cells were grown in the presence of 0.1 mM 3 from OD_{600}=0.2 and T4 lysozyme expression, induced by the addition of 0.2 mM arabinose after 1h. After a further 3 h cells were harvested. Tagged proteins in the lysate were detected via an inverse electron demand Diels-Alder reaction between incorporated **3** and tetrazine fluorophore **4a** (20 mM, 1h, RT).

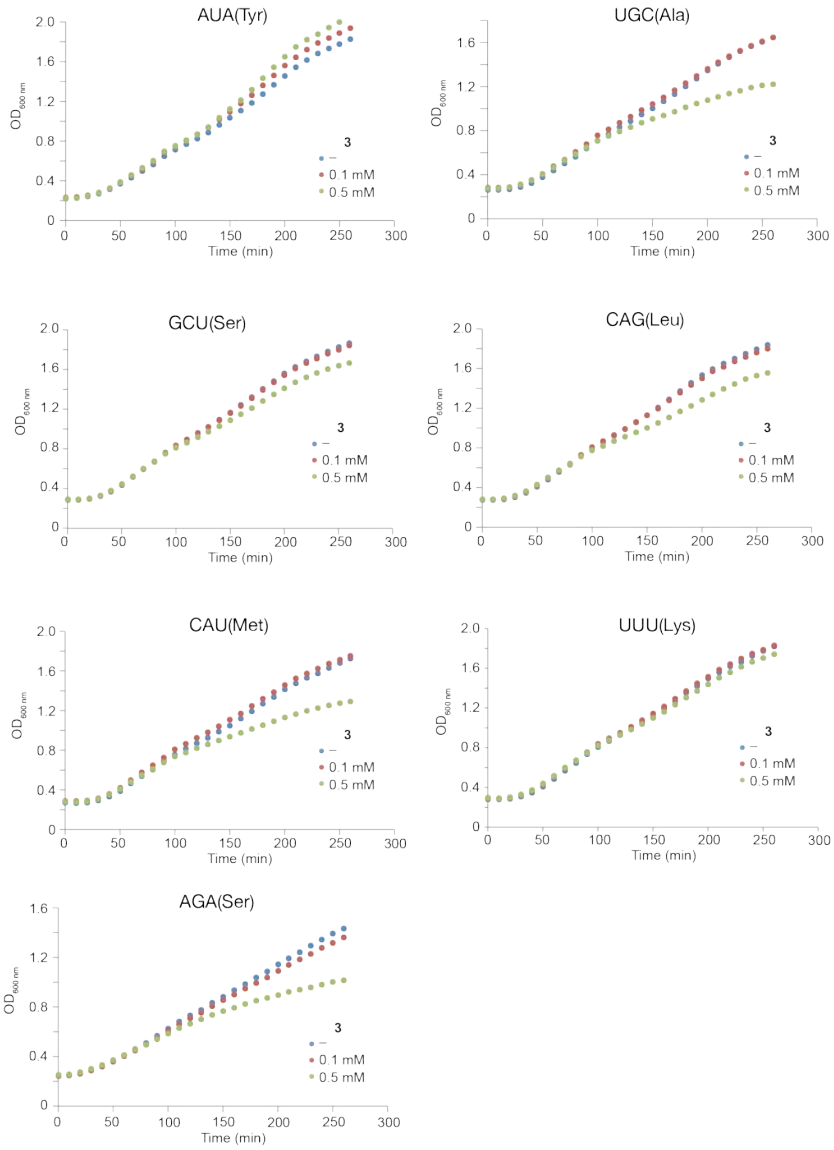
The amino acids in parentheses are the natural amino acids encoded by the endogenous tRNA bearing the corresponding anti-codon. **B**. Lane profile analysis for each codon.

Supplementary Figure 4. Specific amino acid replacement in SORT demonstrated by ESI-MS



T4 lysozyme isolated after SORT with UUU(Lys) in the presence of 1mM **3**. Expected mass WT T4 lysozyme: 19512.2 Da, Found mass: 19510 ± 2.0 Da. Expected mass WT T4 lysozyme Lys→**3** single mutation: 19622.3 Da, Found mass: 19620 ± 2.0 Da.

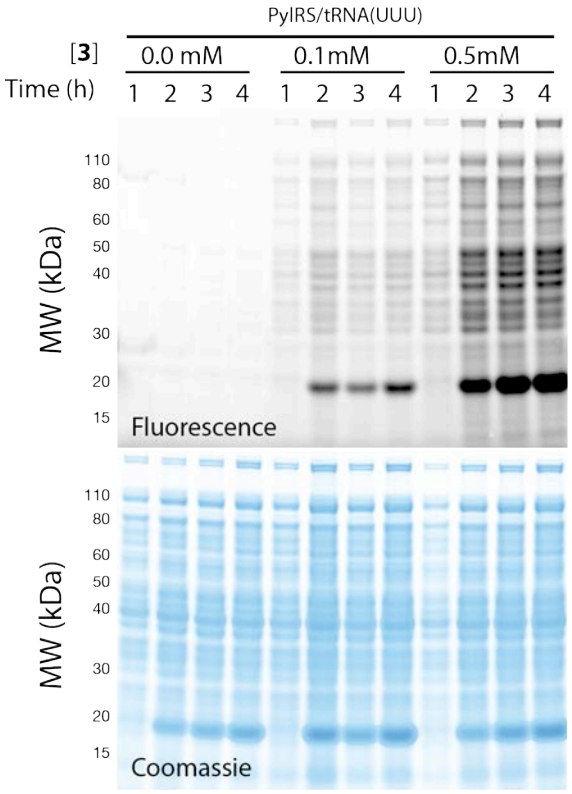
Supplementary Figure 5: Incorporation of 3 (0.1 mM) via SORT-M is not toxic to cells



Chemically competent DH10B cells were transformed with two plasmids: pBKwtPylRS necessary for expression of PylRS, and pBAD_wtT4L_MbPylT$_{XXX}$ plasmids that is required for expression of PyltRNA$_{XXX}$ and expresses lysozyme under arabinose control. The cells were recovered in 1 ml SOB medium for one hour at 37°C prior to aliquoting to 10 ml LB-KT (LB media with 50 μg ml^{-1} kanamycin, and 25 μg ml^{-1} tetracycline) and incubated overnight (37°C, 250 rpm, 12 h). The overnight culture (OD_{600}≈3) was diluted to a OD_{600}~0.3 in 10 mL LB-KT$_{1/2}$ (LB media with 25 μg ml^{-1} kanamycin, and 12.5 μg ml^{-1} tetracycline) supplemented with **3** at different concentrations, 0, 0.1, 0.5 mM. 200 μL aliquots of these cultures were transferred into a 96-well plate and OD_{600} measured using a Microplate reader,

Infinite 200 Pro (TECAN). OD_{600} was measured for each sample every 10 min with linear 1 mm shaking between the measurements.

Supplementary Figure 6: Measurement of time-dependent variation in incorporation of 3 in proteome via SORT-M at different concentrations of 3 in response to AAA codon

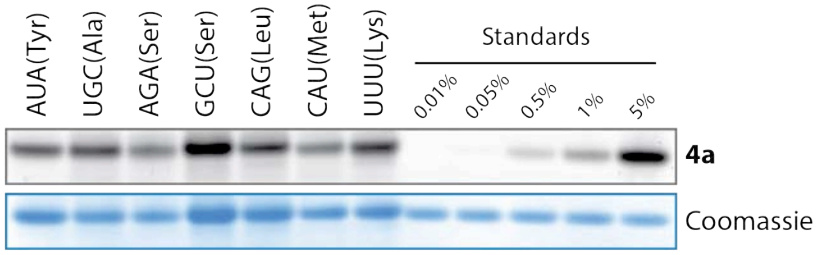


Chemically competent DH10B cells were transformed with two plasmids: pBKwtPylRS necessary for expression of PylRS, and pBAD_wtT4L_MbPylT$_{UUU}$ plasmid that is required for expression of PyltRNA$_{UUU}$. pBAD_wtT4L_MbPylT$_{UUU}$ plasmid also contains the gene for expression of T4 lysozyme that is downstream of arabinose-inducible promoter. After transformation, cells were recovered in 1 ml SOB medium for one hour at 37°C prior to inoculation in 10 ml LB-KT (LB media with 50 μg ml^{-1} kanamycin, and 25 μg ml^{-1} tetracycline). The culture was incubated overnight (37°C, 250 rpm, 12 h) and subsequently diluted to an OD_{600}~0.3 in 30 mL LB-KT$_{1/2}$ (LB media with 25 μg ml^{-1} kanamycin, and 12.5 μg ml^{-1} tetracycline) supplemented with **3** at different concentrations, 0, 0.1, 0.5 mM. The cultures was incubated (37°C, 250 rpm) for 1 h, when OD_{600} reached approximately 0.6. 2 ml culture aliquot was collected in a separate tube for each of three cultures. This is the pre-induction culture (lane labelled as1 in the gel image). Subsequently arabinose was added at a final concentration of 0.2% (v/v) to induce expression of T4 lysozyme and culture aliquots of 2 mL were collected every hour (lanes labelled as 2, 3 and 4 corresponding to 1, 2 and 3h

culture collection after induction). For each of the collected cultures, bacterial cells were pelleted by centrifugation at 4 ºC, washed with ice cold PBS (3 x 1 mL) and subsequently the pellets were frozen and stored at -20 ºC. The pellets were then thawed in 200 µL of ice cold PBS and lysed by sonication (9 x 10 s ON / 20 s OFF, 70% power). The lysates were clarified by centrifugation at 15,000 RPM, 4 ºC for 30 minutes. The supernatants were transferred to fresh 1.5 mL tubes. 50 µL of supernatant was transferred to a new tube for the labeling reactions, and the rest was frozen in liquid nitrogen and stored at -80C. To the 50 µL of supernatant, 0.5 µL of 2 mM **4a** was added and the lysates were incubated at 25ºC for 1 hour. After 1h, 17 µL of 4X LDS sample buffer supplemented (6mM BCN and 5% BME) was added and mixed by vortexing gently. Samples were incubated for 10min before boiling at 90 ºC for 10 min. Samples were analysed by 4-12% SDS-PAGE and fluorescent images were acquired using Typhoon Trio phosphoimager (GE Life Sciences)

Supplementary Figure 7. SORT-M incorporation frequency in *E. coli*

A



B

```
atggatatatttgaaatgttacgtatagatgaaggtcttagacttaaaatctataaagac
 M  D  I  F  E  M  L  R  I  D  E  G  L  R  L  K  I  Y  K  D
acagaaggctattacactattggcatcggtcatttgcttacaaaaagtccatcacttaat
 T  E  G  Y  Y  T  I  G  I  G  H  L  L  T  K  S  P  S  L  N
gctgctaaatctgaattagataaagctattgggcgtaatactaatggtgtaattacaaaa
 A  A  K  S  E  L  D  K  A  I  G  R  N  T  N  G  V  I  T  K
gatgaggctgaaaaactctttaatcaggatgttgatgctgctgttcgcggaattctgaga
 D  E  A  E  K  L  F  N  Q  D  V  D  A  A  V  R  G  I  L  R
aatgctaaattaaaaccggtttatgattctcttgatgcggttcgtcgcgctgcattgatt
 N  A  K  L  K  P  V  Y  D  S  L  D  A  V  R  R  A  A  L  I
aatatggttttccaaatgggagaaaccggtgtggcaggatttactaactctttacgtatg
 N  M  V  F  Q  M  G  E  T  G  V  A  G  F  T  N  S  L  R  M
cttcaacaaaaacgctgggatgaagcagcagttaacttagctaaaagtagatggtataat
 L  Q  Q  K  R  W  D  E  A  A  V  N  L  A  K  S  R  W  Y  N
caaacacctaatcgcgcaaaacgagtcattacaacgtttagaactggcacttgggacgcg
 Q  T  P  N  R  A  K  R  V  I  T  T  F  R  T  G  T  W  D  A
tataaaaatctaagccatcatcatcatcatcattaa
 Y  K  N  L  S  H  H  H  H  H  H  -
```

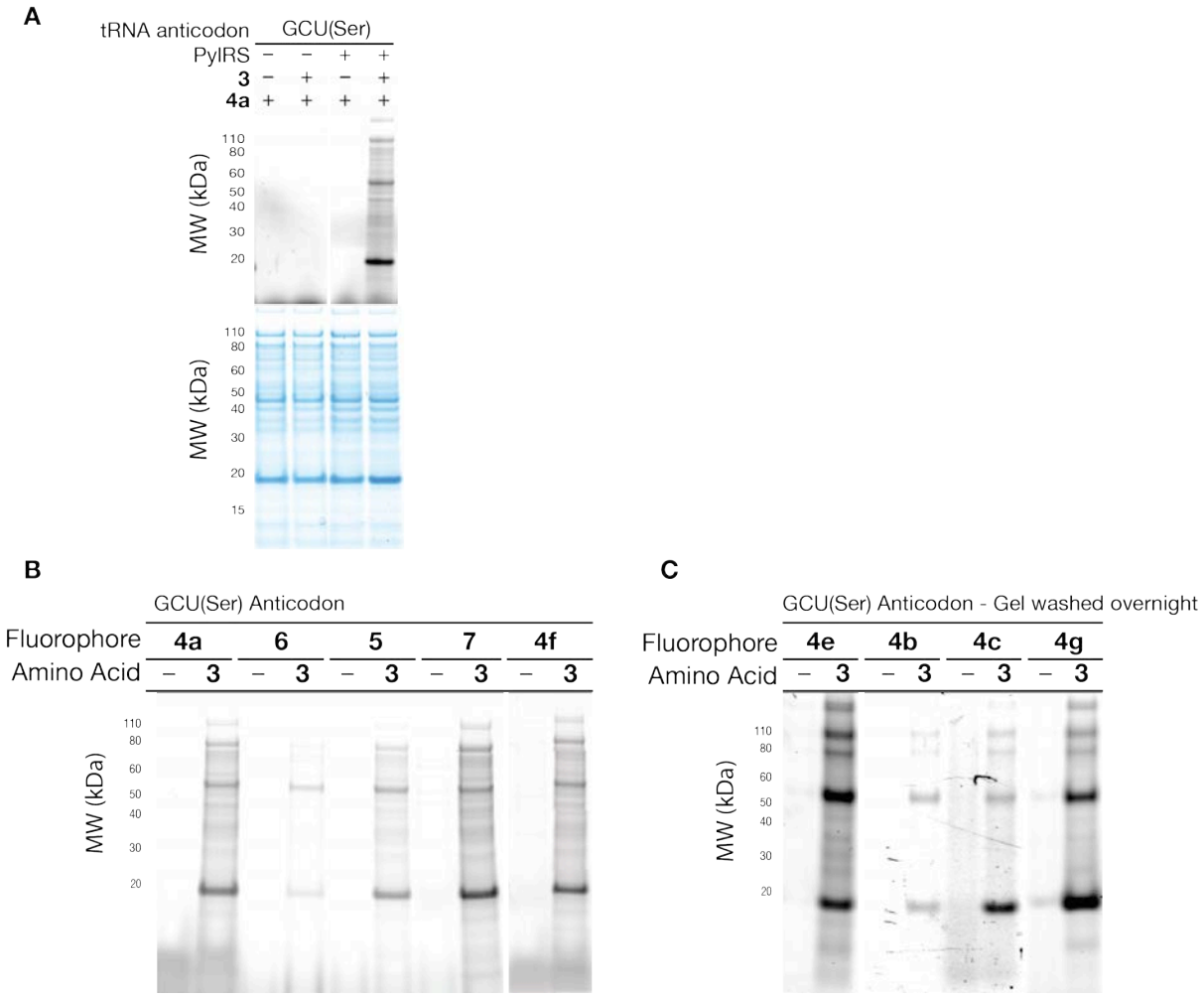
C

tRNA Anticodon	Approximate label count (per 1000 molecules)	Paired codons in T4Lysozyme	% incorporation per codon
AUA(Tyr)	9.19	5	0.18
UGC(Ala)	15.61	5 (7)	0.31 (0.22)
AGA(Ser)	12.05	3	0.40
GCU(Ser)	19.52	1 (3)	1.95 (0.65)
CAG(Leu)	14.47	1	1.45
CAU(Met)	5.54	5	0.11
UUU(Lys)	16.11	13 (21)	0.12 (0.08)

A. To quantify the frequency of incorporation of **3** in to T4 lysozyme via SORT-M, a series of standards were made by mixing quantitatively labelled recombinant T4 lysozyme (T4L K83-**3-4a,** made by genetically encoding 3 in response to an amber codon at position 83 in the T4 lysozyme gene using the PylRS/$tRNA_{CUA}$ pair and quantitatively labelling the resulting protein with **4a**, with unlabelled T4L K83-**3** to create the indicated standards. 1 μg of each standard was loaded along side each *E. coli* lysate labelled using SORT-M (labelled

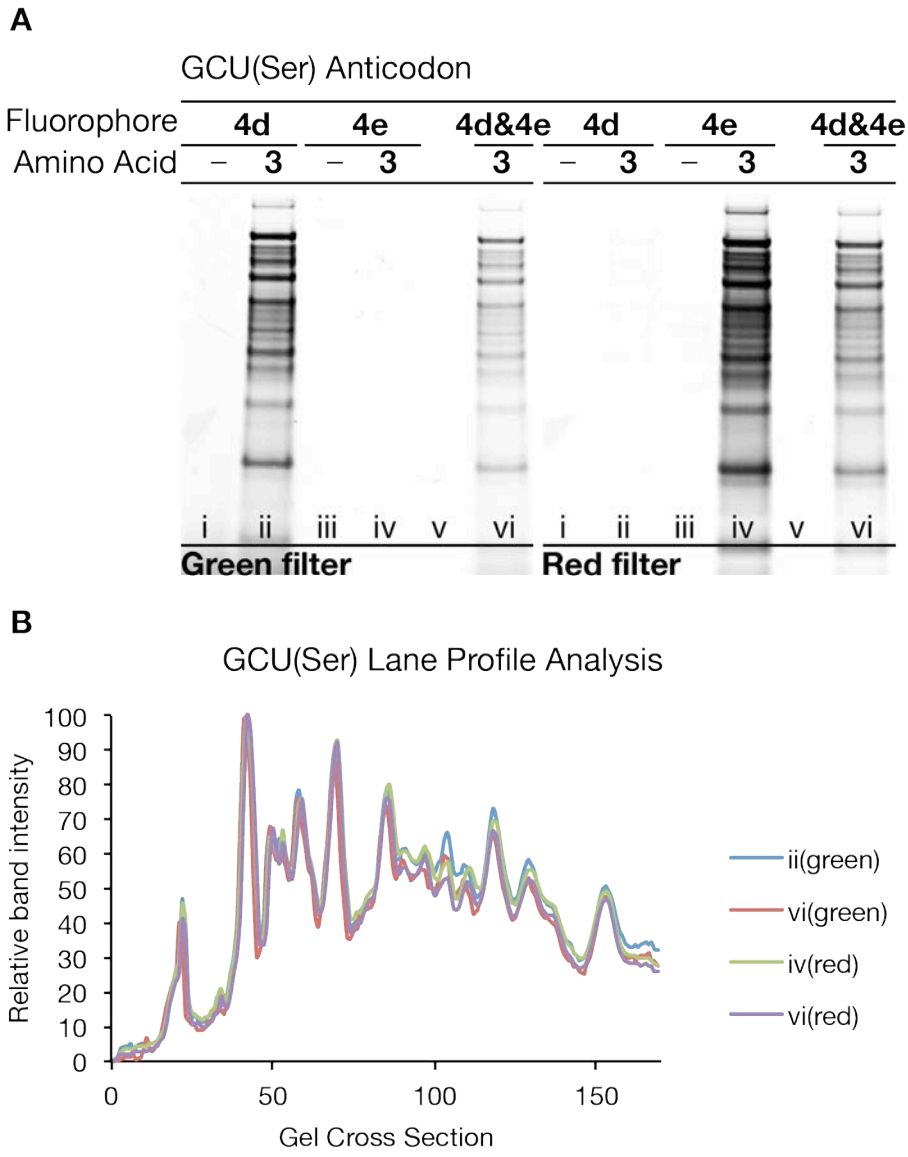
with **4a**). **B.** Sequence of T4 lysozme. **C.** Densitometry measurements of T4 lysozyme labelled with 4a from each SORT-M lysate were normalised by the intensity of the coomassie band and compared against a standard curve generated from the standards. The number of labels per T4 lysozyme was calculated for each SORT-M experiment and the percent incorporation per codon calculated by diving the number of labels per molecule by the number of relevant codons. Calculations were performed for strict Watson-Crick decoding and, where relevant, for wobble decoding- (data shown in brackets).

Supplementary Figure 8. A variety of tetrazines modules and fluorophores can be used for SORT-M in *E. coli*



A. In-gel fluorescence analysis of the *E. coli* proteome with **3** (1mM) incorporated by SORT. **B.** The proteome was labelled with tetrazine-dye conjugates containing different tetrazine modules (please see **Figure 2B** for structures). **C**. Fluorescence analysis of the same proteome as above, this time labelled with tetrazine-dye conjugates containing the same tetrazine but differing conjugated dyes (please see **Figure 2B** for structures, fluorescent images were collected on a Typhoon Trio imaging system (GE Healthcare) for dye **4e** emission filter set to 580 nm and excitation laser set at 532 nm, for dye **4b** emission filter set to 526 nm and excitation laser set at 532 nm, for dye **4c** emission filter set to 526 nm and excitation laser set at 532 nm, for dye **4g** emission filter set to 670 nm and excitation laser set at 633 nm).

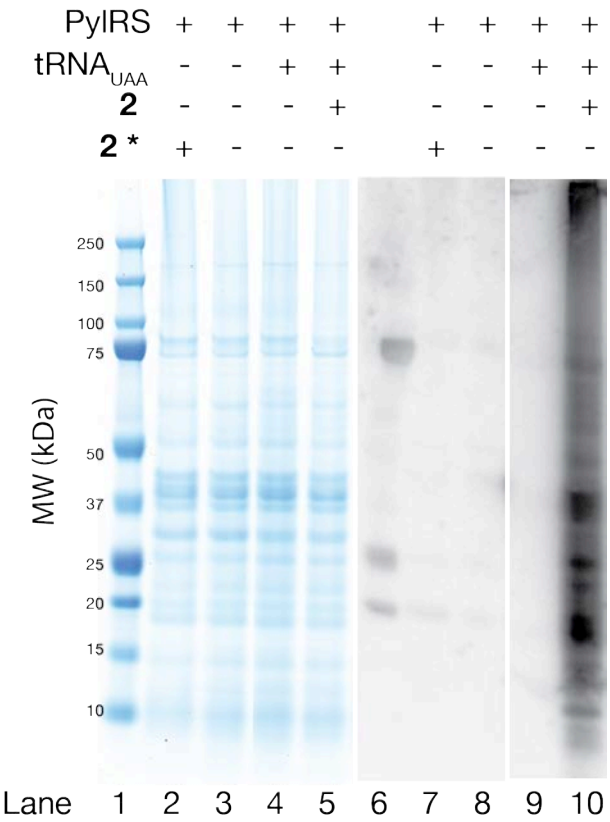
Supplementary Figure 9. SORT-M works comparably with distinct tetrazine fluorophores



A. An *E.coli* proteome was tagged via SORT using **3** and the PylRS/tRNA$_{GCU}$ pair. The lysate was split in two and one half labelled with a red fluorophore **4e (Lane iv)** while the other half was labelled with a green fluorophore **4d (Lane ii)**, fluorescent images were collected on a Typhoon Trio imaging system (GE Healthcare) for dye **4d** emission filter set to 526 nm and excitation laser set at 532 nm, for dye **4e** emission filter set to 580 nm and excitation laser set at 532 nm). The labelling reactions were mixed and loaded in **Lane vi**. Lanes **i-vi** were imaged with a green filter (left side) and red filter (right side). **B.** Lane profile analysis was performed on lanes **ii (**imaged in the green channel), **iv** (imaged in the red channel) and **vi** (imaged in both the green and red channels) the band intensities were then

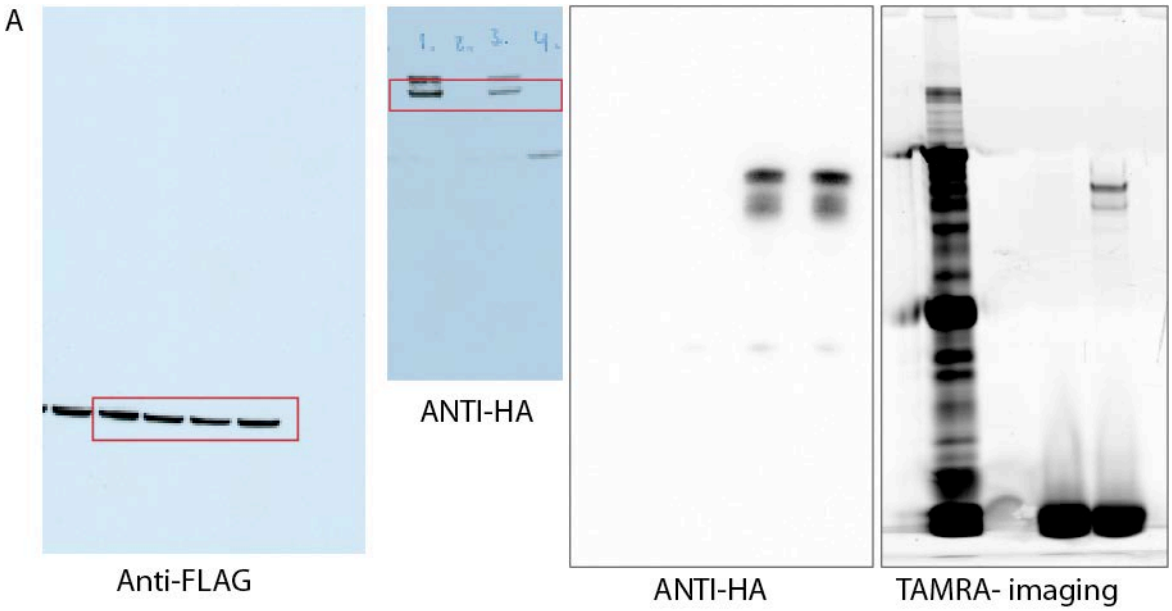
plotted (normalised to the most intense band). The overlay shows very high correlation between the two labelling reactions indicating that the choice of probe does not substantially bias the labelling.

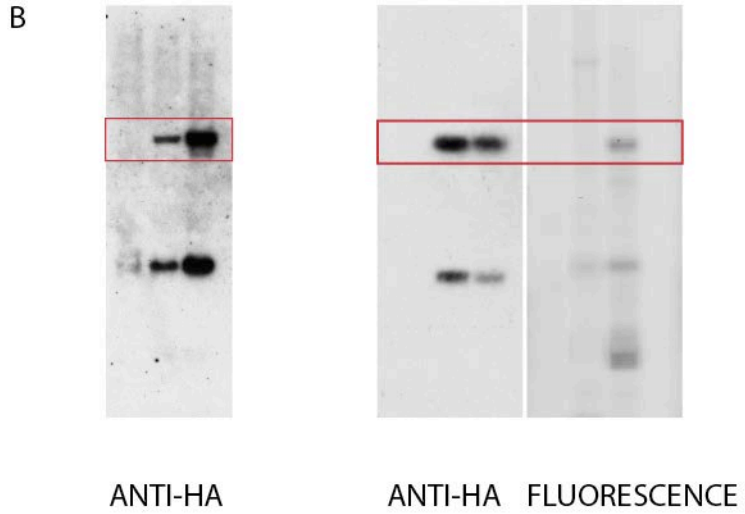
Supplementary Figure 10. SORT-M using 2



E. coli proteome incorporating **2** then labelled with a azide-biotin conjugate and imaged using a streptavidin-HRP conjugate, after depletion of naturally occurring biotinylated proteins. 2* indicates samples where **2** is added to the cell lysate prior to labelling reaction to control for the effects of non-encoded free amino acid on labelling. Proteome labelling via this approach was less efficient than cycloprpene labelling in our hands. Lanes 1 and 6 are the protein ladder.

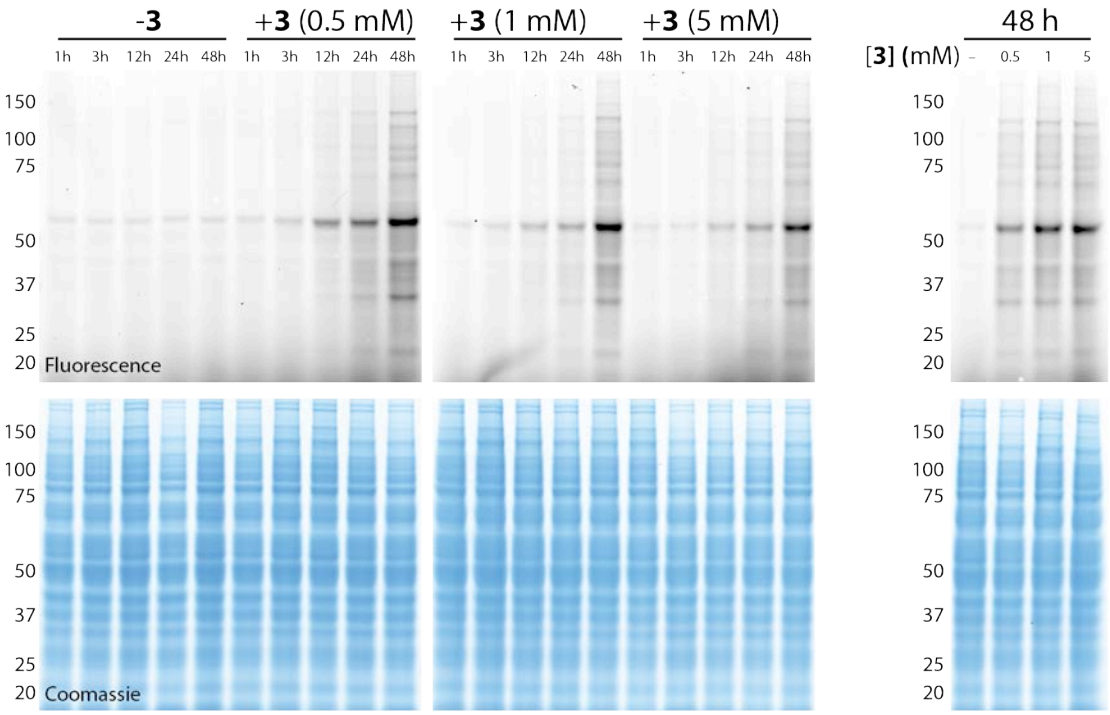
Supplementary Figure 11. Full blots from Figures in main paper





A. Full blots from **Figure 2**. **B**. Full blots from **Figure 3**.

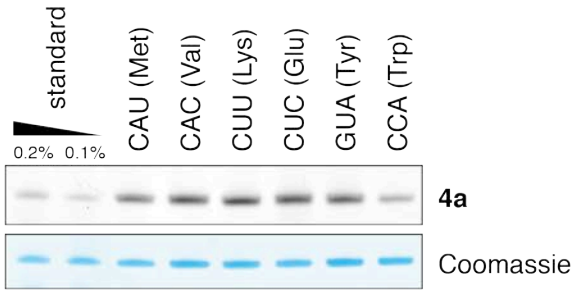
Supplementary Figure 12. Time and concentration dependence of SORT-M labelling in mammalian cells



Cells were plated on 24 well plates and grown to near confluence. The cells were transfected using TurboFect (Thermo Scientific) with the pCMV-MmPylS[1] and pU6-PyltRNA$_{CUC}$ constructs. Cells were grown for 48 hours after transfection. **3** was added to the media to a final concentration of 0.0, 0.5, 1, or 5 mM either 1, 3, 12, 24, or 48 hours before harvesting and lysis in ice cold PBS. The lysates were clarified by centrifugation at 4°C and 21,000 x g, and the tetrazine-dye conjugate **4g** was added to the supernatants to a final concentration of 4 μM. The labeling reaction was incubated at 25°C for 60 minutes. The reactions were then quenched by heating to 95°C in 4X LDS sample buffer (Life technologies) supplemented with BCN to give a final concentration of 1 mM, and analyzed by SDS-PAGE. After destaining in ultrapure water, the gels were imaged on a Typhoon Trio imaging system (GE Healthcare).

Supplementary Figure 13. The frequency of protein labelling by SORT-M in mammalian cells

A



B

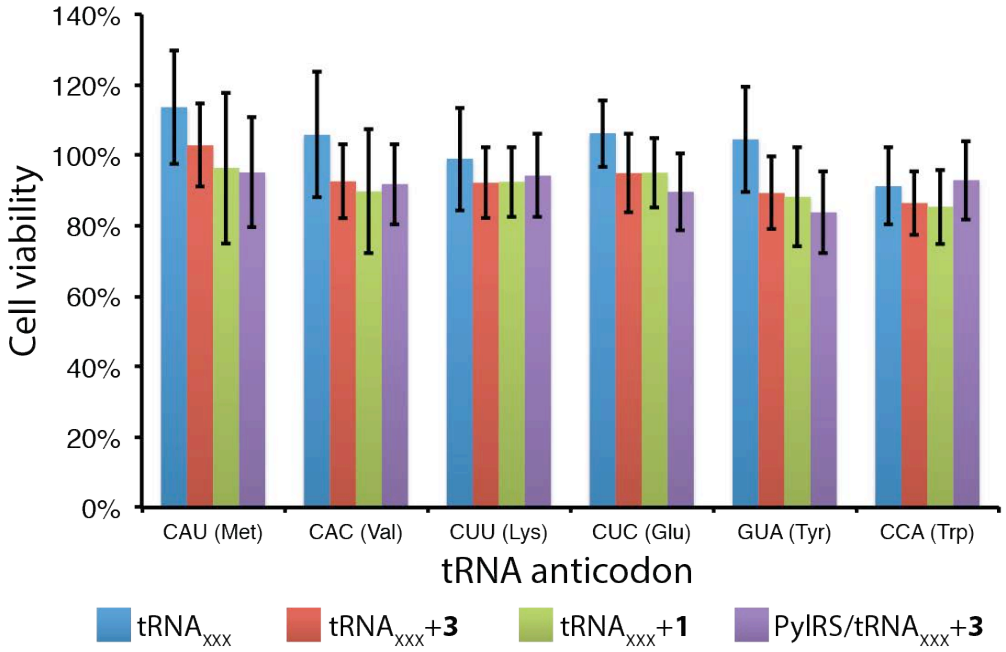
```
atggtgtccaagggcgaggaactgttcaccggcgtggtgcccatcctggtggaactggat
 M  V  S  K  G  E  E  L  F  T  G  V  V  P  I  L  V  E  L  D
ggcgacgtgaacggccacaagttctctgtgcggggagagggcgaaggcgacgccacaaat
 G  D  V  N  G  H  K  F  S  V  R  G  E  G  E  G  D  A  T  N
ggcaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccttggcctacc
 G  K  L  T  L  K  F  I  C  T  T  G  K  L  P  V  P  W  P  T
ctcgtgaccacactgacctacggcgtgcagtgcttcagcagataccccgaccatatgaag
 L  V  T  T  L  T  Y  G  V  Q  C  F  S  R  Y  P  D  H  M  K
cggcacgacttcttcaagagcgccatgcccgagggctacgtgcaggaacggaccatcagc
 R  H  D  F  F  K  S  A  M  P  E  G  Y  V  Q  E  R  T  I  S
ttcaaggacgacggcacctacaagaccagagccgaagtgaagttcgagggcgacaccctc
 F  K  D  D  G  T  Y  K  T  R  A  E  V  K  F  E  G  D  T  L
gtgaaccggatcgagctgaagggcatcgatttcaaagaggacggcaacatcctgggccac
 V  N  R  I  E  L  K  G  I  D  F  K  E  D  G  N  I  L  G  H
aagctggagtacaacttcaacagccacctggtgtacatcaccgccgacaagcagaagaac
 K  L  E  Y  N  F  N  S  H  L  V  Y  I  T  A  D  K  Q  K  N
ggcatcaaggccaacttcaagatccggcacaacgtggaagatggcagcgtgcagctggcc
 G  I  K  A  N  F  K  I  R  H  N  V  E  D  G  S  V  Q  L  A
gaccactaccagcagaacacccccatcggagatggccccgtgctgctgcccgacaaccac
 D  H  Y  Q  Q  N  T  P  I  G  D  G  P  V  L  L  P  D  N  H
tacctgagcacccagagcgtgctgagcaaggaccccaacgagaagcgggaccacatggtg
 Y  L  S  T  Q  S  V  L  S  K  D  P  N  E  K  R  D  H  M  V
ctgctggaattcgtgaccgccgctggcatcacccacggcatggacgagctgtacaagggc
 L  L  E  F  V  T  A  A  G  I  T  H  G  M  D  E  L  Y  K  G
agccaccatcaccatcaccattgataa
 S  H  H  H  H  H  H  -  -
```

C

tRNA anticodon	approximate label count (per 1000 molecules)	Paired codons in sfGFP	% incorporation per codon
CAT (Met)	3.37	5	0.07
CAC (Val)	4.42	19	0.02
CTT (Lys)	4.48	19	0.02
CTC (Glu)	4.80	9	0.05
GTA (Tyr)	4.30	9	0.05
CCA (Trp)	2.00	1	0.2

A. sfGFP coexpressed with the indicated sense PylRS/$tRNA_{XXX}$, pairs with **3** (0.5mM) was purified from a 10 cm dish of HEK293-T cells by anti GFP affinity chromatography. The purified protein (1μg), and standards bearing defined amounts of sfGFP containing **3** were labelled with **4a** (2μM, 4h at RT). The standard is sfGFP that has **3** incorporated at a known frequency. To create this standard sfGFP bearing **3** quantitatively incorporated at position 150 (SfGFP-**3**) was expressed using the PylRS/$tRNA_{CUA}$ pair and an *sfGFP-His6* gene bearing an amber codon at position 150. The resulting protein was mixed in defined ratios with sfGFP-His6 that does not contain any unnatural amino acids to give the indicated percentage of GFP molecules containing **3**. **B**. Amino acid and DNA sequence of SfGFP-His6 and its gene. **C**. Calculated percent incorporation per codon. Densitometry of fluorescent labelling for sense codon incorporation experiments was compared to the standards to provide an incorporation frequency per molecule of GFP. The sequence of GFP was used to define the number of cognate codons for each tRNA. The percent incorporation per codon was calculated by dividing the percent incorporation per GFP by the number of relevant codons in GFP.

Supplementary Figure 14. SORT-M labelling has minimal effect on mammalian cell viability



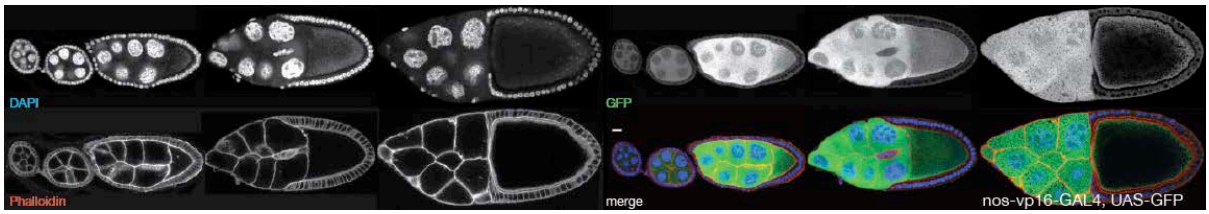
HEK293-T cells were transfected as described in the methods section and grown for 48 hours in DMEM + 10% FBS either without unnatural amino acid, with 0.5 mM **1**, or with 0.5 mM **3**. MTT labelling reagent (3-[4,5-dimethythiazol-2-yl]-2,5-diphenyl tetrazolium bromide) was then added to the culture media for four hours, prior to solubilization of the formazan product with 10% SDS in 10 mM HCl. Cell viability was quantified from the difference in absorbance at 550 nm and 690 nm, and normalized to untransfected controls (100%). Error bars show the standard deviation of three trials.

Supplementary Figure 15. Amino acid and DNA sequence of Drosophila GFP-amber-mCherry-HA

```
atggtgtccaagggcgaggagctgtttaccggcgtggtgcccattctggtggagctggat
 M  V  S  K  G  E  E  L  F  T  G  V  V  P  I  L  V  E  L  D
ggcgacgtgaacggccacaagttcagcgtgtccggcgagggcgagggcgacgccacctat
 G  D  V  N  G  H  K  F  S  V  S  G  E  G  E  G  D  A  T  Y
ggaaagctgaccctgaagttcatctgcaccaccggcaagctgcccgtgccatggccaacc
 G  K  L  T  L  K  F  I  C  T  T  G  K  L  P  V  P  W  P  T
ctcgtgaccaccctgacctatggcgtgcagtgcttcagccgctaccccgatcacatgaag
 L  V  T  T  L  T  Y  G  V  Q  C  F  S  R  Y  P  D  H  M  K
cagcacgatttcttcaagagcgccatgcccgagggctacgtgcaggagcgcaccatcttt
 Q  H  D  F  F  K  S  A  M  P  E  G  Y  V  Q  E  R  T  I  F
ttcaaggatgacggcaactacaagacccgcgccgaagtgaagttcgagggcgataccctc
 F  K  D  D  G  N  Y  K  T  R  A  E  V  K  F  E  G  D  T  L
gtgaaccgcatcgagctgaagggcatcgatttcaaggaggatggaaacatcctgggccac
 V  N  R  I  E  L  K  G  I  D  F  K  E  D  G  N  I  L  G  H
aagctggagtacaactacaacagccacaacgtgtacatcatggccgacaagcagaagaac
 K  L  E  Y  N  Y  N  S  H  N  V  Y  I  M  A  D  K  Q  K  N
ggcatcaaagtgaacttcaagatccgccacaacatcgaggatggcagcgtgcagctggcc
 G  I  K  V  N  F  K  I  R  H  N  I  E  D  G  S  V  Q  L  A
gatcactaccagcagaacaccccaatcggcgacggcccagtgctgctgcccgataaccat
 D  H  Y  Q  Q  N  T  P  I  G  D  G  P  V  L  L  P  D  N  H
tacctgagcacccagagcgccctgagcaaggaccccaacgagaagcgcgaccacatggtg
 Y  L  S  T  Q  S  A  L  S  K  D  P  N  E  K  R  D  H  M  V
ctgctggagtttgtgaccgccgccggcattaccctgggcatggatgagctgtacaagcgc
 L  L  E  F  V  T  A  A  G  I  T  L  G  M  D  E  L  Y  K  R
gcccagatccacgatctggtgtaggtgggaggatcgacccgcgtgtcgaagggcgaggag
 A  Q  I  H  D  L  V  -  V  G  G  S  T  R  V  S  K  G  E  E
gataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctccgtg
 D  N  M  A  I  I  K  E  F  M  R  F  K  V  H  M  E  G  S  V
aatggacacgagttcgagattgagggcgagggcgagggacgcccatatgagggaacccag
 N  G  H  E  F  E  I  E  G  E  G  E  G  R  P  Y  E  G  T  Q
accgccaagctgaaagtgaccaagggcggacccctgcccttcgcctgggatattctgagc
 T  A  K  L  K  V  T  K  G  G  P  L  P  F  A  W  D  I  L  S
ccccagtttatgtacggcagcaaggcctacgtgaagcaccccgccgatatccccgattac
 P  Q  F  M  Y  G  S  K  A  Y  V  K  H  P  A  D  I  P  D  Y
ctgaagctgagcttcccagagggcttcaagtgggagcgcgtgatgaatttcgaggacggc
 L  K  L  S  F  P  E  G  F  K  W  E  R  V  M  N  F  E  D  G
ggagtcgtgaccgtgacccaggatagcagtttgcaggatggcgagttcatctacaaagtg
 G  V  V  T  V  T  Q  D  S  S  L  Q  D  G  E  F  I  Y  K  V
aagctgcgcggcaccaacttcccgtccgatggcccagtgatgcagaagaagaccatgggc
 K  L  R  G  T  N  F  P  S  D  G  P  V  M  Q  K  K  T  M  G
tgggaggccagcagcgagcgcatgtatccagaggatggcgccctgaagggcgagatcaag
 W  E  A  S  S  E  R  M  Y  P  E  D  G  A  L  K  G  E  I  K
cagcgcctgaagctgaaggatggcggccactacgatgccgaagtgaagaccacctacaag
 Q  R  L  K  L  K  D  G  G  H  Y  D  A  E  V  K  T  T  Y  K
gccaagaagccggtgcagctgccaggcgcctacaatgtgaacatcaagctggatatcacc
 A  K  K  P  V  Q  L  P  G  A  Y  N  V  N  I  K  L  D  I  T
tcccacaacgaggactacaccatcgtggagcagtatgagcgcgccgagggccgccatagt
 S  H  N  E  D  Y  T  I  V  E  Q  Y  E  R  A  E  G  R  H  S
accggcggaatggacgagctgtataagatgtaccctacgatgtgcccgattacgccgag
 T  G  G  M  D  E  L  Y  K  M  Y  P  Y  D  V  P  D  Y  A  E
cagaagctgatctccgaggaggacctgcaccatcaccaccaccacggaagtggcagcggc
 Q  K  L  I  S  E  E  D  L  H  H  H  H  H  H  G  S  G  S  G
tccccaaagaagaagcgcaaggtgtaa
 S  P  K  K  K  R  K  V  -
```

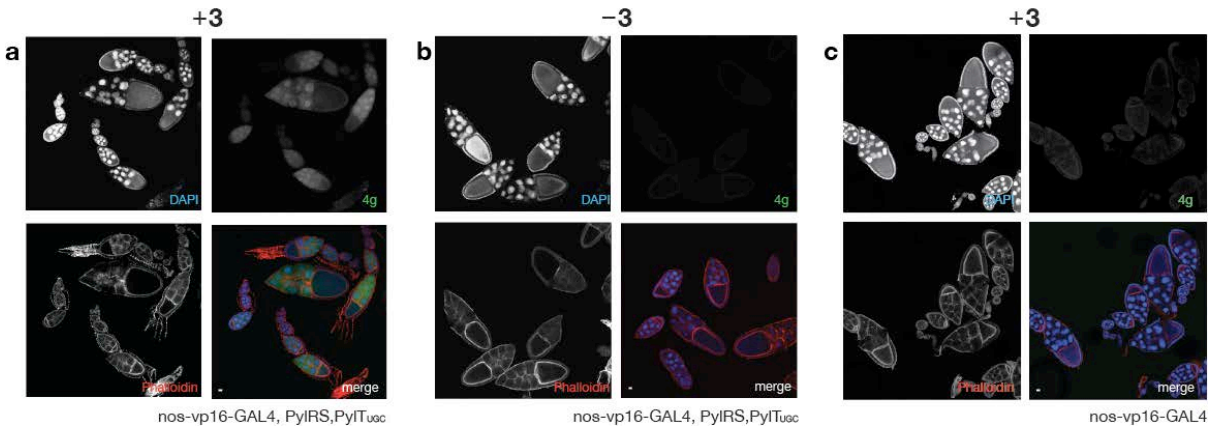
GFP (amino acid residues 1-238), Amber codon at position 248, mCherry (amino acid residues 255-489), HA tag (amino acid residues 491-499), Myc tag (amino acid residues 500-509), His tag (amino acid residues 510-515) and SV40 NLS (amino acid residues 523-528).

Supplementary Figure 16. Specificity of nos-vp16-GAL4 driven GFP expression within the ovary.



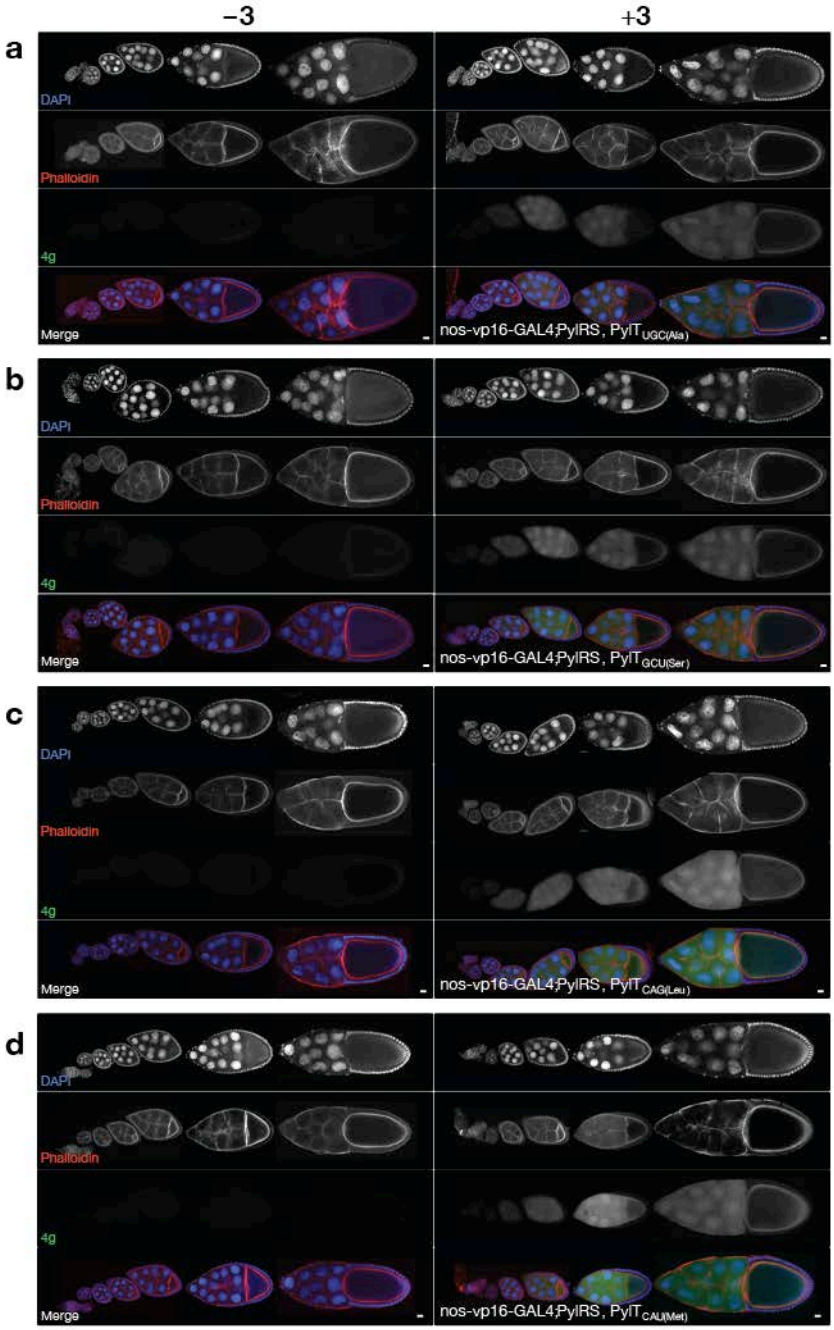
Egg chambers expressing UAS-GFP under the control of nos-vp16-GAL4 show progressive GFP accumulation specifically in germ line cells starting from stage 5 to stage 10. Fly egg chamber from stages 4 to 10 of oogenesis (anterior to the left). Ovaries from flies of the indicated genotypes were dissected, fixed and stained using DAPI (blue) and Alexa-594 conjugated Phalloidin (red). Scale bar is 20 μm.

Supplementary Figure 17. SORT-M works reproducibly in a population of individuals of the same genotype



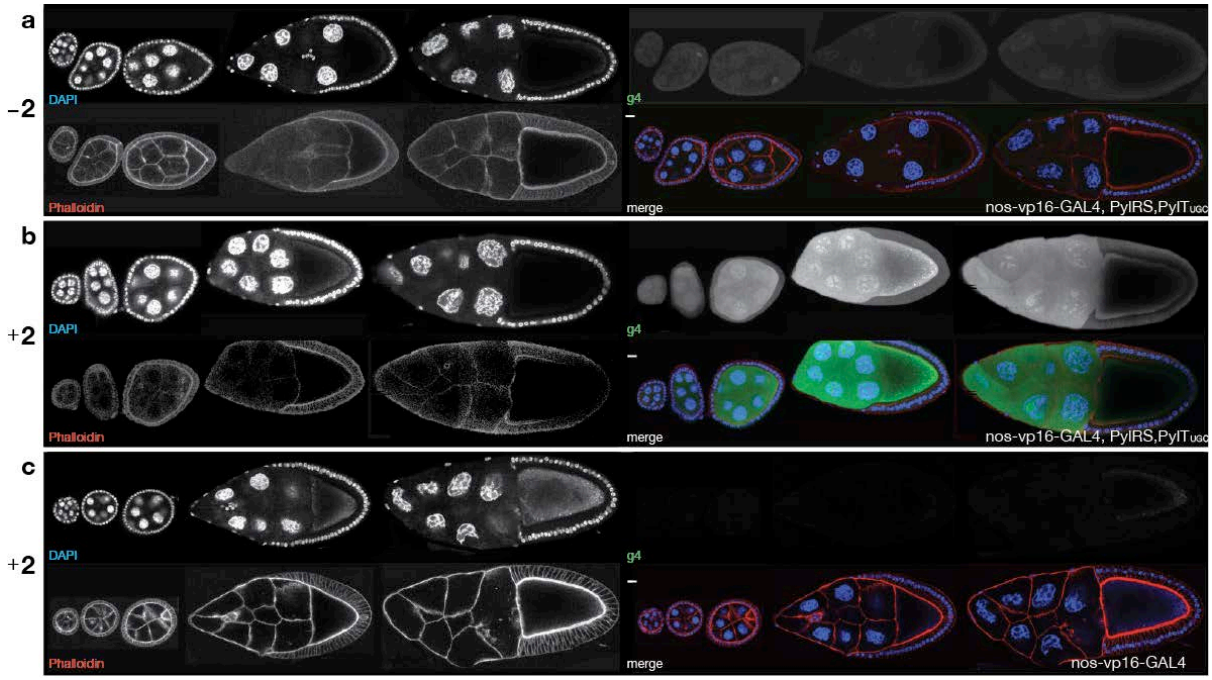
A-C. Wide-field images of ovary labelling samples of the indicated genotype and feeding condition. Stained with DAPI (blue), Phalloidin (red) and **4g** (green). Scale bar is 20 μm.

Supplementary Figure 18. Cell-type and temporally specific SORT-M in flies using different anticodons



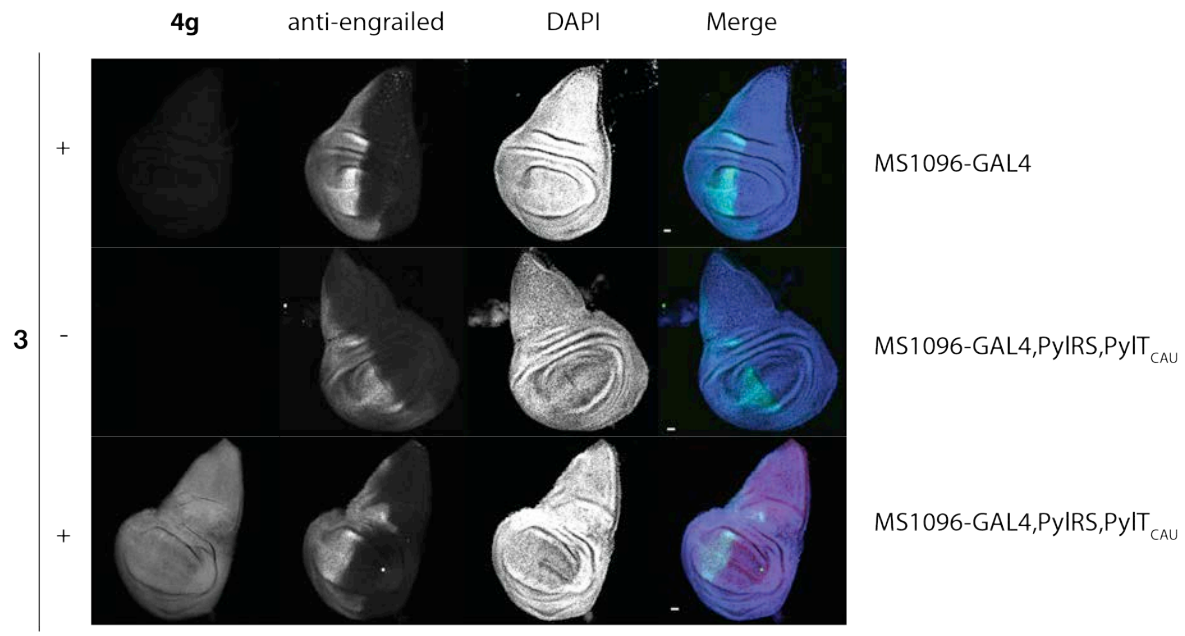
A-D. Genetically directed, cell type specific proteome labelling using the indicated anticodons. Egg chambers from stage 4 to stage 10 of oogenesis of the indicated genotypes and feeding condition are shown. Anterior to the left, DAPI (blue), Phalloidin (red), 4g (green). Scale bar is 20 μm.

Supplementary Figure 19. Cell-specific SORT-M using 2



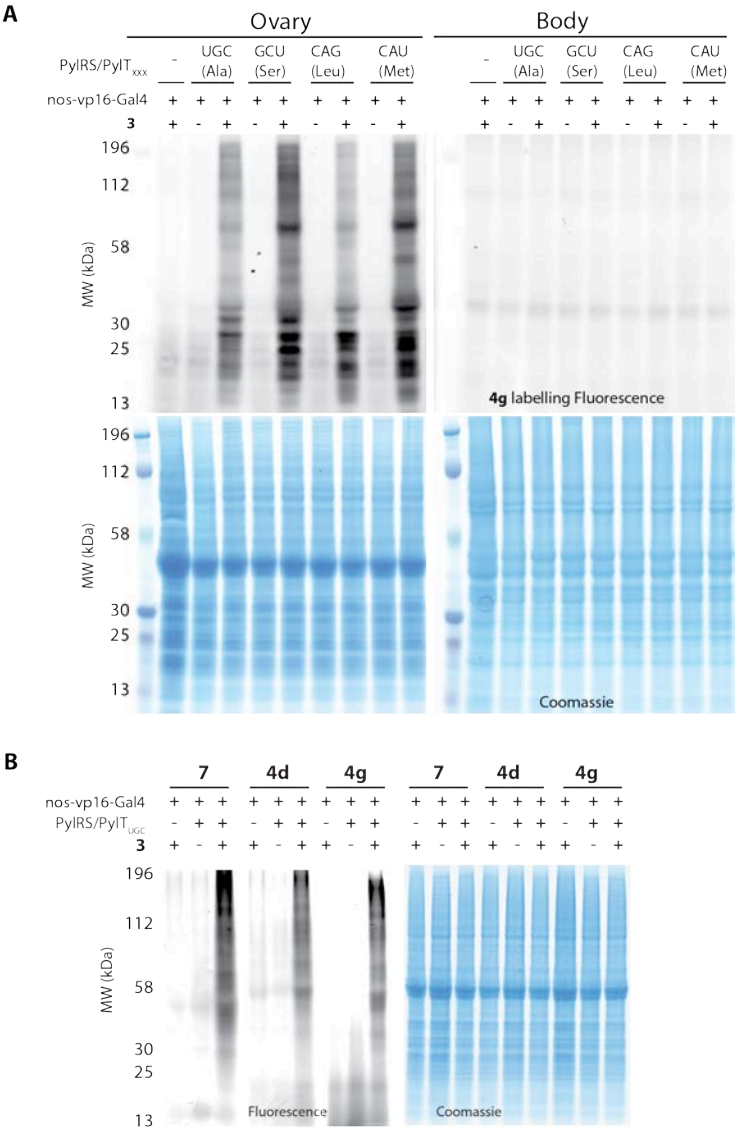
A-C. Genetically directed, cell type specific proteome labelling within an organ. Egg chambers from stage 4 to stage 10 of oogenesis of the indicated genotypes and feeding condition are shown. Anterior to the left, DAPI (blue), Phalloidin (red), Alexa 488-azide (green), scalebar 20μm. **A**. Egg chambers not receiving **2**, expressing PylRS/PylT$_{UGC}$ and labelled with Alexa 488-azide**. B**. Egg chambers receiving **2**, expressing PylRS/PylT$_{UGC}$ and labelled with Alexa 488-azide**. C**. Egg chambers receiving **2**, not expressing PylRS/PylT$_{UGC}$ and labelled with Alexa 488-azide**.** Scale bar is 20 μm.

Supplementary Figure 20. Amino acid dependent proteome labelling of the wing imaginal disc



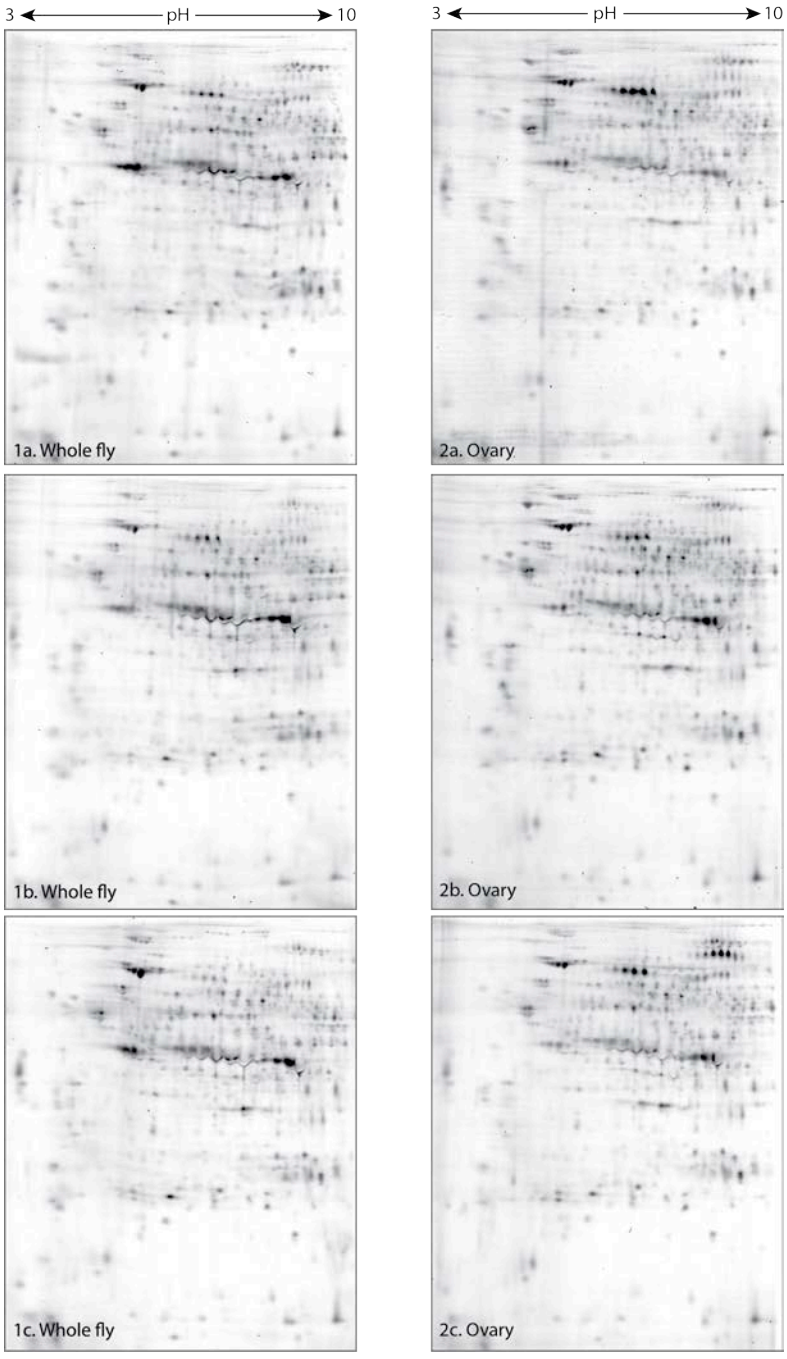
Wing imaginal discs were dissected from 3rd instar larvae of a cross between MS1096-GAL4 females and PylRS, PylT$_{CAU}$ males that were grown on medium in the presence or absence of **3**. Specific proteome labelling with **4g** is only observed in the wing discs of larvae that express PylRS, PylT$_{CAU}$ and are fed with **3**. Scale bar is 20 μm.

Supplementary Figure 21. Ovary specific proteome labelling via SORT-M



A. Labeling of proteins extracted from the ovaries and bodies of nos-vp16-GAL4 or nos-vp16-GAL4, PylRS, PylT$_{xxx}$ flies fed normal food or food supplemented with **3**. Labelling is dependent on feeding **3** and occurs selectively in ovary extracts, but not in the rest of the body samples. **B**. Ovary proteome labelling with different tetrazine dye conjugates. Comparable labelling results were obtained with different tetrazine conjugated dyes (**7**, **4d** and **4g** from **Figure 1**) indicating that different dyes do not influence labelling specificity in fly lysates. The top panels in each figure are fluorescent images from a Typhoon phosphoimager.

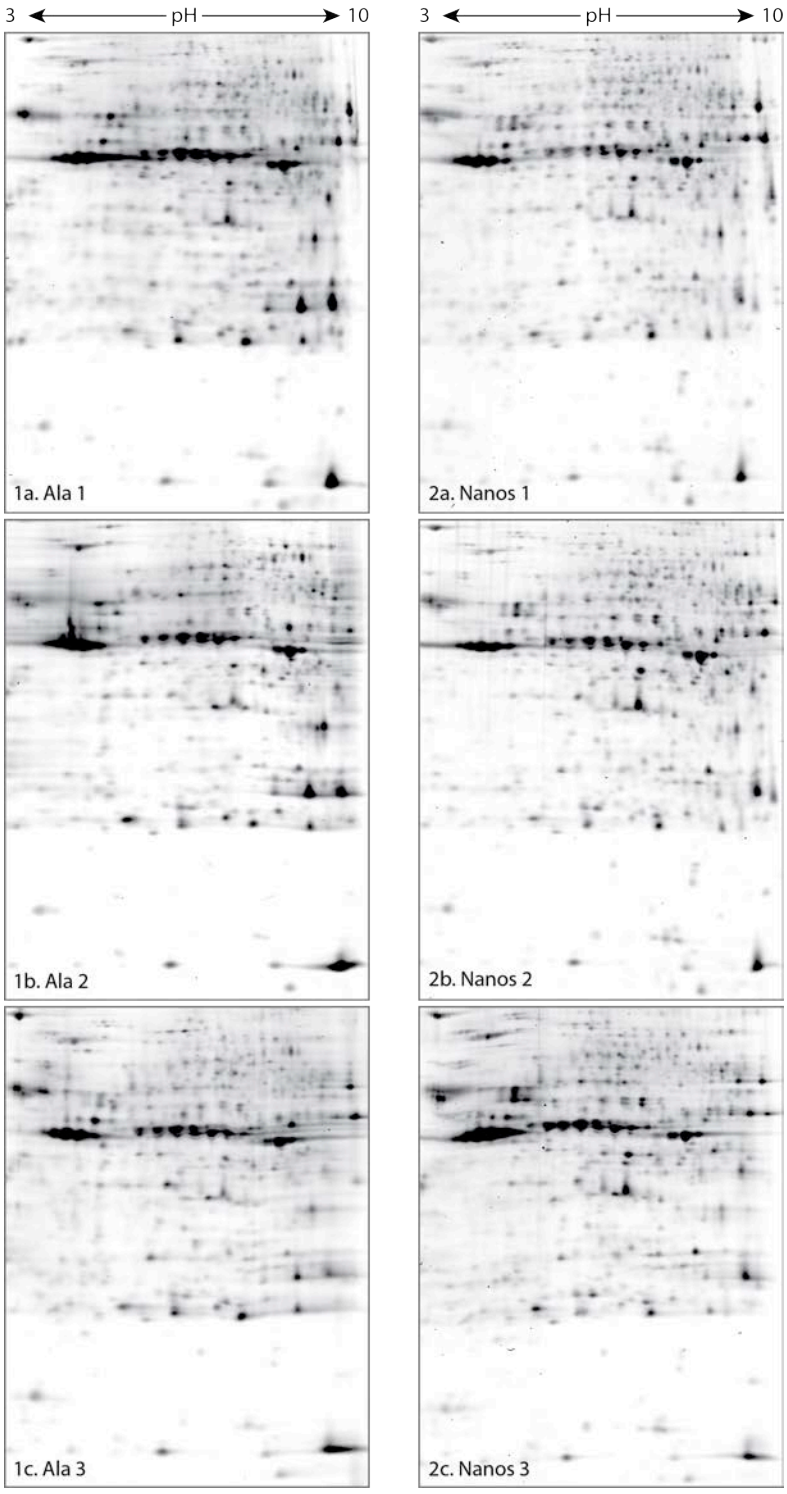
Supplementary Figure 22. Comparing ovary and body labelling with 4e



nos-vp16-GAL4, PylRS, $PylT_{UGC}$ flies fed **3,** and dissected ovaries from these flies were both labelled with **4e.** Bodies plus ovaries or ovaries alone from three independent collections were labelled with **4e** and subjected 2D-DIGE. The selective identification of ovary proteins without dissection requires the labeling of whole nos-vp16-

GAL4,PylRS,PylT_{UGC} flies fed **3** with **4e** and the selective detection of ovary proteins (**Fig. 6d**). We quantified the similarity of proteome labeling with **4e** on whole flies to the labeling with **4e** on dissected ovaries by comparing three replicates of ovary labelling (isolated from nos-vp16-GAL4, PylRS,PylT_{UGC} flies fed **3)** to three replicates of labeling whole nos-vp16-GAL4, PylRS,PylT_{UGC} flies fed **3** by 2D-DIGE. Biological variation analysis (**Online Methods**) revealed that only 4.0% of spots (from a total of 649 spots) showed a 2.3-fold difference in spot intensity ($p<0.05$, power >0.8). Taking into account false discovery rates led to only 2 spots (0.3%) that remain significantly different between the two labellings (q value<0.05). These results demonstrate the specificity of the method for selectively labeling a genetically targeted tissue in the presence of the proteome derived from non-targeted tissue.

Supplementary Figure 23. Three biological replicates of Cy3-NHS ester stained 2D-DIGE for nos-vp16-GAL4, PylRS, $PylT_{UGC}$ flies fed with amino acid 3 and three biological replicates of Cy3-NHSester stained 2D-DIGE for nos-vp16-GAL4 flies



The proteomes were extracted from 3 biological replicates of nos-vp16-GAL4, PylRS, $PylT_{UGC}$ flies fed normal food supplemented with **3** and 3 biological replicates of the nos-

vp16-GAL4 flies fed un-supplemented normal food (50 flies per replicate). An internal standard was created containing equal amounts of protein (100μg) from each of the six samples (600 μg total fly protein) and labeled with Cy5-NHS ester according to the manufacturers instructions (Minimal Labelling Kit , Cyanagen). In a parallel labelling reaction, 50 μg of each biological replicate was labelled with Cy3-NHS ester. The Cy3 labelling of each biological replicate was mixed with 50 μg of the internal standard and subject to 2D-DIGE with isoelectric focusing in the first dimension using 13 cm IPG strips (pH 3-10, non-linear), followed by a second dimension separation by SDS-PAGE. The gels were imaged using a Typhoon scanner to visualize and quantify the Cy3 and Cy5 fluorescence. "Nanos" on plots refers to nos-vp16-Gal4 flies, while "Ala" on plots refers to nos-vp16-GAL4, PylRS, $PylT_{UGC}$ flies fed **3**. We used the biological variation analysis module within the Decyder software package for 2D-DIGE analysis to perform paired student t-tests on gel spots[2-4]. Only 2.5% of spots (from 600 spots) showed a 2.3 -fold difference in spot intensity ($p<0.05$, power >0.8). Upon taking into account the false discovery rate in the large data sets, through the application of Storey's q-value analysis as previously described for 2D-DIGE data sets[2-4], only three protein spots (0.3%) out of those previously identified remained significantly different (q-value <0.05). These experiments demonstrate that the proteome of nos-vp16-GAL4 flies is very similar to the proteome of nos-vp16-GAL4,PylRS,$PylT_{UGC}$) flies fed **3**, and that the expression of the PylRS, $PylT_{UGC}$ pair and labeling of flies with **3** does not significantly affect their proteome, as judged by 2D-DIGE and biological variation analysis.[2-4]

Supplementary Figure 24. Silver stain images of 2D-gel electrophoresis experiments for MS/MS analysis and protein identification

A. Fly ovary proteins were first labelled with **4e**, then mixed with fly body extracts from the same flies and analysed by 2D-gel electrophoresis. **B**. Fly ovary protein extracts were first mixed with fly body extracts then labelled with **4e** and analysed by 2D-gel electrophoresis. **C**. Fly ovary protein extracts were labelled with **4e** (and not mixed with fly body extracts) and then analysed by 2D-gel electrophoresis. nos-vp16-GAL4, PylRS, $PylT_{UGC}$ flies fed **3** were used for all experiments. In each case the spots circled in **Figure 6,** were identified by staining with **4e** and the gel was silver stained to allow spot picking for MS/MS analysis.

Supplementary Table 1. Primers

Name	Sequence	Notes
FMT17	CATGTAGATCGAATGGACTATAAATCCG TTCAGCCGGG	Mutagenesis primers for D.m. PylT
FMT18	CCCGGCTGAACGGATTTATAGTCCATTC GATCTACATG	CTA to ATA (Tyr)
FMT19	CATGTAGATCGAATGGACTTGCAATCCG TTCAGCCGGGTTAG	Mutagenesis primers for D.m. PylT
FMT20	CTAACCCGGCTGAACGGATTGCAAGTCC ATTCGATCTACATG	CTA to TGC (Ala)
FMT21	GTAGATCGAATGGACTAGAAATCCGTTC AGCCGGG	Mutagenesis primers for D.m. PylT
FMT22	CCCGGCTGAACGGATTTCTAGTCCATTC GATCTAC	CTA to AGA (Ser)
FMT23	CATGTAGATCGAATGGACTGCTAATCCG TTCAGCCGGGTTAG	Mutagenesis primers for D.m. PylT
FMT24	CTAACCCGGCTGAACGGATTAGCAGTCC ATTCGATCTACATG	CTA to GCT (Ser)
FMT25	GTAGATCGAATGGACTTAAAATCCGTTC AGCCGGG	Mutagenesis primers for D.m. PylT
FMT26	CCCGGCTGAACGGATTTTAAGTCCATTC GATCTAC	CTA to TAA (Leu)
FMT27	GTAGATCGAATGGACTCAGAATCCGTTC AGCCGGG	Mutagenesis primers for D.m. PylT
FMT28	CCCGGCTGAACGGATTCTGAGTCCATTC GATCTAC	CTA to CAG (Leu)
FMT29	GTAGATCGAATGGACTCATAATCCGTTC AGCCGGG	Mutagenesis primers for D.m. PylT
FMT30	CCCGGCTGAACGGATTATGAGTCCATTC GATCTAC	CTA to CAT (Met)
FMT31	CATGTAGATCGAATGGACTTTTAATCCG TTCAGCCGGGTTAG	Mutagenesis primers for D.m. PylT
FMT32	CTAACCCGGCTGAACGGATTAAAAGTCC ATTCGATCTACATG	CTA to TTT (Lys)
MAM01	GTAGATCGAATGGACTCATAATCCGTTC AGCCGGG	Forward mutagenesis primer for CTA to CAT (Met)
MAM02	CCCGGCTGAACGGATTATGAGTCCATTC GATCTAC	Reverse mutagenesis primer for CTA to CAT (Met)

MAM03	GATCGAATGGACTCACAATCCGTTCAGCCGG	Forward mutagenesis primer for CTA to CAC (Val)
MAM04	CCGGCTGAACGGATTGTGAGTCCATTCGATC	Reverse mutagenesis primer for CTA to CAC (Val)
MAM05	GATCGAATGGACTCTTAATCCGTTCAGCCG	Forward mutagenesis primer for CTA to CTT (Lys)
MAM06	CGGCTGAACGGATTAAGAGTCCATTCGATC	Reverse mutagenesis primer for CTA to CTT (Lys)
MAM07	GATCGAATGGACTCTCAATCCGTTCAGCCG	Forward mutagenesis primer for CTA to CTC (Glu)
MAM08	CGGCTGAACGGATTGAGAGTCCATTCGATC	Reverse mutagenesis primer for CTA to CTC (Glu)
MAM09	GTAGATCGAATGGACTGTAAATCCGTTCAGCCG	Forward mutagenesis primer for CTA to GTA (Tyr)
MAM10	CGGCTGAACGGATTTACAGTCCATTCGATCTAC	Reverse mutagenesis primer for CTA to GTA (Tyr)
MAM11	GATCGAATGGACTCCAAATCCGTTCAGCCG	Forward mutagenesis primer for CTA to CCA (Trp)
MAM12	CGGCTGAACGGATTTGGAGTCCATTCGATC	Reverse mutagenesis primer for CTA to CCA (Trp)

Supplementary Table 2. Hatching rates

Fly Line	Amino Acid 3	Hatching Rate (%)	Sample (n)
y w	-	90.9	452
y w	+	92.3	644
UAS-GFP	-	84.1	435
UAS-GFP	+	81.1	355
UGC (Ala)	-	86.1	274
UGC (Ala)	+	94.1	324
GCU (Ser)	-	96.9	388
GCU (Ser)	+	93.4	412
CAG (Leu)	-	68.7	498
CAG (Leu)	+	66.7	580
CAU (Met)	-	93.0	415
CAU (Met)	+	95.0	545

The hatching rates for nos-vp16-GAL4, PylRS,$PylT_{UGC}$ flies are 86.1% (-**3**, n=274), and 94.1% (+**3**, n=324). Comparison to hatching rates for yellow/white "wild-type" embryos[5]: 90.9% (-**3**, n=452) and 92.3% (+**3**, n=644), or nos-vp16-GAL4/UAS-GFP: 84.1% (-**3**, n=435) and 81.1 % (+**3**, n=355), demonstrates that proteome labelling does not have a negative effect on hatching. Egg chamber development also proceeds normally for the additional anticodons tested and led to hatching rates as follows: GCU(Ser), (-**3**, 96.9%, n=388, +**3**, 93.4%, n= 412); CAG (Leu), (-**3**, 68.7%, n=498, +**3**, 66.7%, n= 580); CAU (Met), (-**3**, 93.0%, n=415, +**3**, 95.0%, n= 545). These data further confirm that SORT-M labeling of serine and methionine codons, like SORT-M at alanine codons has very little effect on hatching rate. The CAG (Leu) hatching rate also shows very little amino acid dependence, suggesting that

proteome labeling with **3** is not problematic in this line. However, the hatching rate for this line is lower than that for the other synthetase/tRNA pairs and this may simply result from insertion of the synthetase or tRNA genes at an unfavorable site(s) in the genome.

Supplementary Table 3. MS/MS identification of proteins excised from 2D-gels 1, 2 and 3

				Gel1			Gel2			Gel3			modENCODE RNA-Seq	
Spot number	**Identified protein**	**Accession Number**	**MW (kDa)**	Unique peptides	Spectral counts	Coverage	Unique peptides	Spectral counts	Coverage	Unique peptides	Spectral counts	Coverage	Ovary, 4 day female	Carcass, 4 day adult
1	Ubiquitin carboxyl-terminal hydrolase OS=Drosophila melanogaster GN=CG12082 PE=2 SV=1	Q9VZU7_DROME	92	26	41	33%	28	51	40%	33	47	47%	61	12
2,3,4,5	Transitional endoplasmic reticulum ATPase TER94 OS=Drosophila melanogaster GN=TER94 PE=1 SV=1	TERA_DROME	89	81	31	46%	67	26	43%	71	35	60%	210	46
2,3,4,5	Hsc70Cb, isoform A OS=Drosophila melanogaster GN=Hsc70Cb PE=2 SV=1	Q9VUC1_DROME	89	16	21	23%	22	27	37%	12	17	18%	186	63
6, 7	CG5355 OS=Drosophila melanogaster GN=CG5355 PE=2 SV=2	Q9VKW5_DROME	86	18	23	23%	21	23	34%	28	42	42%	127	30
6, 7	Dipeptidyl peptidase 3 OS=Drosophila melanogaster GN=DppIII PE=2 SV=2	DPP3_DROME	89	7	8	9%	10	11	14%	12	12	19%	162	30
8, 9, 10, 11, 12, 13	Heat shock protein 83 OS=Drosophila melanogaster GN=Hsp83 PE=1 SV=1	HSP83_DROME	82	13	15	20%	15	15	23%	18	24	26%	2107	947
8, 9, 10, 11, 12, 13	Heat shock 70 kDa protein cognate 4 OS=Drosophila melanogaster GN=Hsc70-4 PE=1 SV=3	HSP7D_DROME	71	7	7	13%	12	14	22%	10	12	19%	2180	1506
13	Vasa, isoform C OS=Drosophila melanogaster GN=vas PE=4 SV=1	M9PBB5_DROME	72	4	4	6%	4	5	6%	5	6	7%	42	4
14, 15, 16	CG8258 OS=Drosophila melanogaster GN=CG8258 PE=2 SV=1	Q7K3J0_DROME	59	15	18	29%	15	16	28%	5	5	10%	150	27
14, 15, 16	Serine/threonine-protein phosphatase PP2A 65 kDa regulatory subunit OS=Drosophila melanogaster GN=Pp2A-29B PE=2 SV=4	2AAA_DROME	65	5	5	9%	8	10	14%	5	7	8%	132	42
17, 18	CG7033 OS=Drosophila melanogaster GN=CG7033 PE=2 SV=2	Q9W392_DROME	58	18	33	36%	24	62	58%	15	27	33%	361	41
17, 18	Inositol-3-phosphate synthase OS=Drosophila melanogaster GN=Inos PE=1 SV=1	INO1_DROME	62	4	4	7%	9	12	21%	5	6	10%	30	31

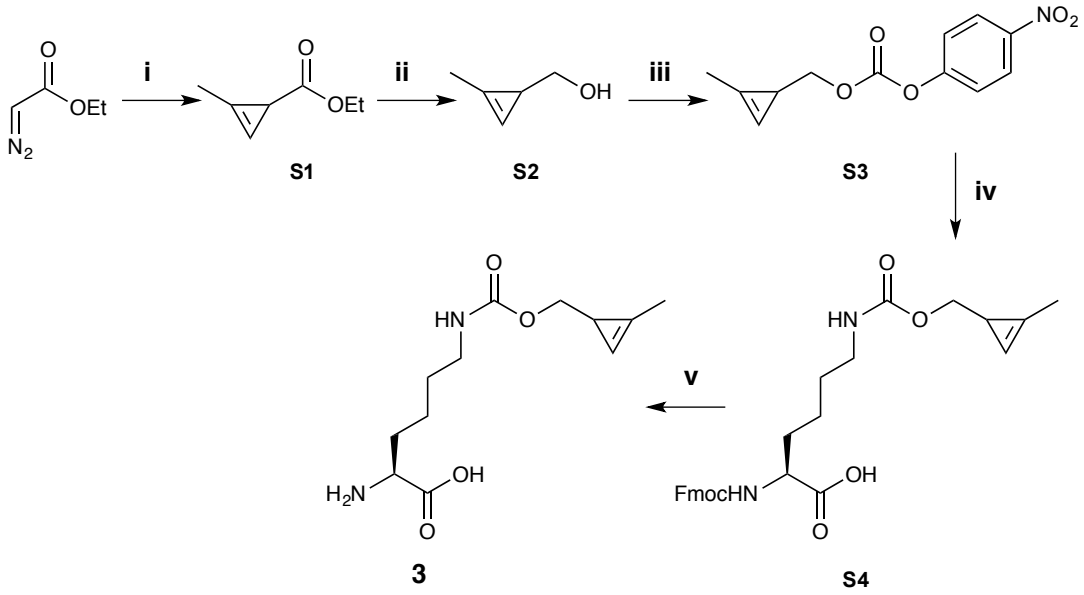
20, 21	T-complex chaperonin 5, isoform B OS=Drosophila melanogaster GN=Cct5 PE=3 SV=1	A1Z8U4_DROME	56	18	49	43%	23	75	57%	18	38	48%	261	53
20, 21	T-complex protein 1 subunit alpha OS=Drosophila melanogaster GN=T-cp1 PE=3 SV=1	A4V391_DROME	60	12	14	24%	18	31	39%	8	10	16%	151	21
25, 26, 27	LD47396p OS=Drosophila melanogaster GN=Tcp-1eta PE=2 SV=2	Q9VHL2_DROME	59	27	72	54%	26	83	54%	23	55	48%	164	27
25, 26, 27	Hsp70/Hsp90 organizing protein homolog OS=Drosophila melanogaster GN=Hop PE=2 SV=1	Q9VPN5_DROME	56	16	22	33%	11	18	24%	9	10	20%	308	59
25, 26, 27	Coronin OS=Drosophila melanogaster GN=coro PE=2 SV=1	Q7JVY0_DROME	57	12	14	26%	8	13	16%	6	7	13%	79	42
25, 26, 27	Kinesin light chain OS=Drosophila melanogaster GN=Klc PE=1 SV=1	KLC_DROME	58	10	11	19%	4	5	7%	4	4	7%	66	22
26, 27	EH domain containing protein OS=Drosophila melanogaster GN=Past1 PE=2 SV=1	Q8T6I0_DROME	61	5	5	10%	7	7	13%	6	6	11%	58	37
22, 23	GH13725p OS=Drosophila melanogaster GN=Tcp-1zeta PE=2 SV=1	Q9VXQ5_DROME	58	24	92	37%	24	113	40%	23	72	38%	184	27
22, 28	Cctgamma, isoform B OS=Drosophila melanogaster GN=Cctgamma PE=3 SV=1	A4V303_DROME	59	21	29	39%	22	40	41%	6	6	12%	174	26
23, 28	Actin-interacting protein 1 OS=Drosophila melanogaster GN=flr PE=2 SV=1	WDR1_DROME	67	15	21	26%	21	46	46%	19	28	43%	78	20
28	CG7461 OS=Drosophila melanogaster GN=CG7461 PE=3 SV=2	A1ZBJ2_DROME	68	5	5	8%	7	7	11%	5	5	8%	125	114
24	CG8036, isoform B OS=Drosophila melanogaster GN=CG8036 PE=2 SV=3	Q9VHN7_DROME	68	18	65	29%	3	4	5%	23	40	40%	383	283
29, 30	Aldehyde dehydrogenase OS=Drosophila melanogaster GN=Aldh PE=2 SV=1	A9J7N9_DROME	57	17	35	32%	16	50	32%	16	27	31%	125	226
32, 33, 34	Vitellogenin-2 OS=Drosophila melanogaster GN=Yp2 PE=1 SV=2	VIT2_DROME	50	8	11	25%	5	10	18%	8	12	21%	1680	1160 20
35, 36, 38	Peroxiredoxin 1 OS=Drosophila melanogaster GN=Jafrac1 PE=1 SV=1	PRDX1_DROME	22	10	25	50%	10	35	50%	9	18	55%	347	85
36	Proteasome subunit beta type OS=Drosophila melanogaster GN=Prosbeta1 PE=2 SV=1	A0AQH0_DROME	24	8	10	57%	7	10	52%	6	7	42%	103	43
37	CG18815, isoform D OS=Drosophila melanogaster GN=CG18815-RA PE=2 SV=1	C0PDF4_DROME	24	6	7	37%	9	31	53%	8	11	51%	64	54
38	Proteasome subunit alpha type-2 OS=Drosophila melanogaster GN=Pros25 PE=1 SV=1	PSA2_DROME	26	5	11	20%	9	20	47%	8	25	50%	186	89
31	Eukaryotic initiation factor 4a, isoform E OS=Drosophila melanogaster GN=eIF-4a-RB PE=2 SV=1	C9QP42_DROME	46	19	69	39%	9	17	25%	30	181	82%	607	474

31	CG3731, isoform A OS=Drosophila melanogaster GN=CG3731 PE=2 SV=2	Q9VFF0_DROME	52	15	20	37%	15	20	37%	12	20	35%	387	600

Supplementary Note 1

Synthesis of unnatural amino acids and probes

Synthesis of methylcycloprop-2-en-1-yl}methoxy)carbonyl]-L-lysine (3)



Scheme 1. Synthesis of N^{ε}-[({2-methylcycloprop-2-en-1-yl}methoxy)carbonyl]-L-lysine **3**. *Reagents and conditions*: **i.** $Rh_2(OAc)_4$, propyne, CH_2Cl_2, 4 °C to RT, 75% yield; **ii.** DIBAL-H, CH_2Cl_2, 0 °C to RT; **iii.** 4-nitrophenyl chloroformate, Hünig's base, CH_2Cl_2, RT, 73% yield; **iv.** Fmoc-Lys-OH, Hünig's base, THF/DMF, 4 °C to RT, 82% yield; **v.** NaOH, THF/H_2O, RT, 68% yield.

i. Ethyl 2-methylcycloprop-2-ene-1-carboxylate S1

A 100 mL 2-neck round bottom flask was charged with CH_2Cl_2 (2 mL) and rhodium acetate (442 mg, 1 mmol, 0.05 eq), and fitted with a dry ice condenser. Propyne (approx. 10 mL) was condensed into the rhodium acetate suspension and the flask lowered into a water bath (20 ºC), a steady reflux of propyne was obtained. Ethyl diazoacetate (2.1 mL, 20 mmol, 1eq) was added to the stirred propyne solution drop-wise over 1 h using a syringe pump. The reaction was stirred at room temperature for a further 10 minutes whereby TLC analysis showed the reaction to be complete by after this time. The cyclopropene product was then purified by silica gel flash column chromatography eluting with pentane and diethyl ether (90:10). This gave the desired product **S1** as a colourless volatile liquid (1.9 g, 75% yield). ^{1}H NMR analysis δ_H (400 MHz, $CDCl_3$) 6.35 (1H, t, *J* 1.4), 4.18-4.09 (2H, m), 2.16 (3H, d, *J* 1.3), 2.12 (1H, d, *J* 1.6), 1.26 (3H, t, *J* 7.1); LRMS *m/z* (ES^+) 127.2 $[M+H]^+$.

These values are in good agreement with literature.{Liao, 2004 #1}

ii. and **iii. (2-Methylcycloprop-2-en-1-yl)methyl (4-nitrophenyl) carbonate S3**

DIBAL-H (22.5 mL of a 1M solution in CH_2Cl_2, 22.5 mmol, 1.5 eq) was added drop-wise to a stirred solution of cyclopropene ester **S1** (1.9 g, 15 mmol, 1 eq) in CH_2Cl_2 (15 mL) at -10 ºC. The reaction was stirred at -10 ºC for 20 minutes before quenching with the cautious addition of H_2O (1 mL), then NaOH (1 mL of a 1 M solution in H_2O) and H_2O (2.3 mL). The mixture was stirred for a further 2h at room temperature before it was dried (Na_2SO_4) and filtered. Hunig's base (3.9 mL, 22.5 mmol, 1.5eq) was added to the filtrate (containing crude cyclopropene alcohol **S2**) followed by the addition of 4-nitrophenyl chloroformate (3.3 g, 16.5 mmol, 1.1 eq). After stirring at room temperature for 18 hours a significant colourless precipitate formed, and TLC analysis showed complete consumption of the crude cyclopropene alcohol **S2**. The reaction was diluted with CH_2Cl_2 and then dry loaded onto silica gel, whereby the activated carbonate **S3** was purified by silica gel column chromatography eluting with ethyl acetate and hexane (20:80). This gave the desired cyclopropene carbonate **S3** as a colourless oil (2.7 g, 73% yield over 2 steps). ^{1}H NMR analysis δ_H (400 MHz, $CDCl_3$) 8.28 (2H, d, *J* 9.2), 7.39 (2H, d, *J* 9.2), 6.62 (1H, s), 4.21 (1H, dd, *J* 10.9, 5.3), 4.14 (1H, dd, *J* 10.9, 5.3), 2.18 (3H, d, *J* 1.3), 1.78 (1H, td, *J* 5.3, 1.3).

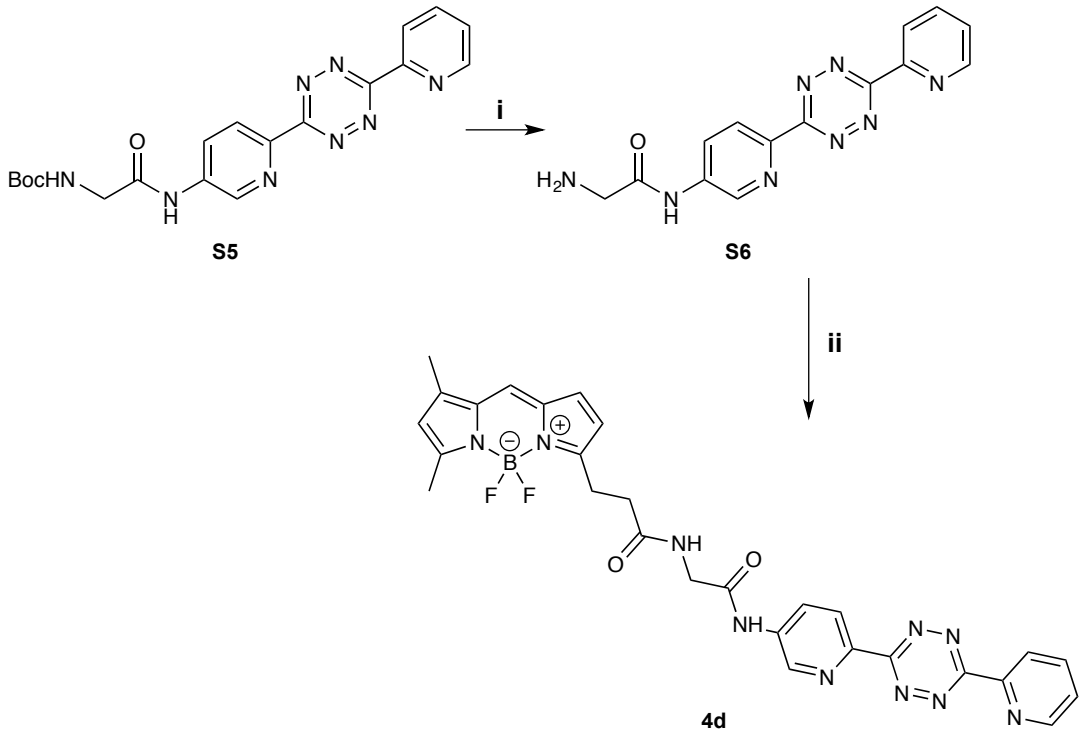
iv. N^{α}-(Fmoc)-N^{ε}-(((2-methylcycloprop-2-en-1-yl)methoxy)carbonyl)-L-lysine S4

Fmoc-Lys-OH·HCl (6.7 g, 16.5 mmol, 1.5 eq) was dissolved in THF (30 mL) and DMF (10 mL), to this solution was added Hünig's base (9.0 mL, 55.0 mmol, 5 eq) followed by cyclopropene carbonate **S3** (2.7 g, 11.0 mmol, 1 eq) an immediate yellow coloration was observed upon addition of the carbonate. The reaction was stirred at room temperature for 6 hours and was adjudged complete by the consumption of starting material after this time as shown by TLC analysis. The crude reaction mixture was dry loaded onto silica gel and the major product purified by silica gel column chromatography eluting with ethyl acetate, hexane and acetic acid (50:49:1 then 99:0:1). This gave the desired product **S4** as a colourless gum (4.3 g, 82% yield). ^{1}H NMR analysis δ_H (400 MHz, $CDCl_3$) 7.77 (2H, t, *J* 7.6), 7.65-7.55 (2H, m), 7.39 (2H, t, *J* 7.6), 7.31 (2H, t, *J* 7.3), 6.54 (1H, s), 5.68-5.57 (1H, m), 4.84 (1H, br-s), 4.44-4.32 (2H, m), 4.22 (1H, t, *J* 7.0), 3.98-3.87 (1H, m), 3.17-3.09 (2H, m), 2.15-2.06 (6H, m), 1.99-1.86 (1H, m), 1.84-1.70 (1H, m), 1.68-1.59 (1H, m), 1.58-1.34 (2H, m); LRMS *m/z* (ES^+) 479.3 $[M+H]^+$, 501.3 $[M+Na]^+$, *m/z* (ES^-) 477.2 $[M-H]^-$.

v. N^{ε}-[({2-methylcycloprop-2-en-1-yl}methoxy)carbonyl]-L-lysine 3

N^{α}-(Fmoc)-N^{ε}-(((2-methylcycloprop-2-en-1-yl)methoxy)carbonyl)-L-lysine **S4** (3.5 g, 7.0 mmol, 1 eq) was dissolved in THF and H_2O (3:1 40 mL), to this solution was added sodium hydroxide (0.9 g, 22.6 mmol, 3.1 eq). The reaction was stirred at room temperature for 8 hours after which time the reaction was adjudged complete by LC-MS analysis. The reaction mixture was diluted with H_2O (100 mL) and the pH adjusted to ~5 by the addition of HCl (1M). The aqueous solution was washed with Et_2O (5× 100 mL), then concentrated to dryness yielding a colourless solid. The solid was purified by preparative HPLC, the product fractions were combined and the solvent removed by freeze-drying. This gave N^{ε}-[({2-methylcycloprop-2-en-1-yl}methoxy)carbonyl]-L-lysine **3** as a colourless solid. δ_H (400 MHz, D_2O) 6.45 (1H, s), 3.90-3.61 (2H, m), 3.09 (1H, t, *J* 6.4), 2.98-2.86 (2H, m), 1.92 (3H, s), 1.52-1.37 (2H, m), 1.37-1.22 (2H, m), 1.21-1.08 (2H, m), 0.83 (1H, d, *J* 5.2). LRMS *m/z* (ES^+) 257.2 $[M+H]^+$, *m/z* (ES^-) 255.2 $[M-H]^-$. δ_C (100 MHz, D_2O) 101.1 (CH), 72.3 (CH_2), 55.9 (CH), 40.2 (CH_2), 34.3 (CH_2), 28.9 (CH_2), 20.3 (CH_2), 16.6 (CH_3), 10.8 (CH) HRMS (ES^+) Found: $(M+Na)^+$ 279.1302. $C_{12}H_{20}O_4N_2Na$ required M^+, 279.1315.

Synthesis of Tetrazine-BODIPY FL 4d



Scheme 3. Synthesis of tetrazine-biotin **4d**. *Reagents and conditions*: **i.** HCl dioxane, RT, 100% yield; **ii.** Bodipy-FL-NHS ester, Hünig's base, DMF, RT.

i. S6

Boc-protected Tetrazine **S6** was synthesized using the procedure reported earlier[6]. 4M HCl in dioxane (500 µL, 2.0 mmol) was added to a stirring solution of Tetrazine **S5** (8 mg, 0.02 mmol) in DCM (500 µL). The reaction was carried out for 2 h at room temperature and 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.

ii. 4d

BODIPY FL succinimidyl ester (5mg, 0.013 mmol, Life technologies) and Hünig's base (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. The reaction

mixture was diluted with 4ml of water and 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; bufferB: acetonitrile). The identity and purity of the tetrazine-BODIPY FL conjugate **4d** was confirmed by LC-MS. ESI-MS: $[M-H]^-$, calcd. 581.38, found 581.2.

Supplementary References

1. Gautier, A. et al. Genetically Encoded Photocontrol of Protein Localization in Mammalian Cells. *Journal of the American Chemical Society* **132**, 4086-4088 (2010).
2. Karp, N.A., Kreil, D.P. & Lilley, K.S. Determining a significant change in protein expression with DeCyder during a pair-wise comparison using two-dimensional difference gel electrophoresis. *Proteomics* **4**, 1421-1432 (2004).
3. Karp, N.A. & Lilley, K.S. Design and analysis issues in quantitative proteomics studies. *Proteomics* **7 Suppl 1**, 42-50 (2007).
4. Lilley, K.S. in Current Protocols in Protein Science (John Wiley & Sons, Inc., 2001).
5. Von Stetina, J.R., Lafever, K.S., Rubin, M. & Drummond-Barbosa, D. A Genetic Screen for Dominant Enhancers of the Cell-Cycle Regulator alpha-Endosulfine Identifies Matrimony as a Strong Functional Interactor in Drosophila. *G3 (Bethesda)* **1**, 607-613 (2011).
6. Lang, K. et al. Genetically encoded norbornene directs site-specific cellular protein labelling via a rapid bioorthogonal reaction. *Nat Chem* **4**, 298-304 (2012).