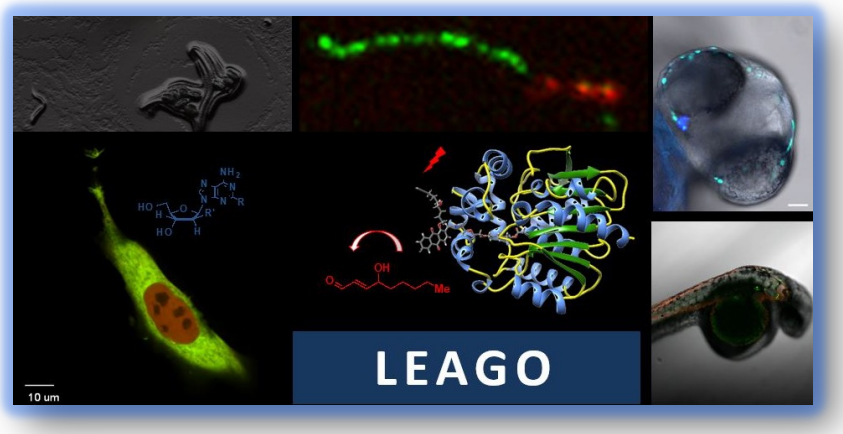
Welcome to CH-313: Chemical Biology

Prof. Yimon Aye https://leago.epfl.ch/



Laboratory of Electrophiles And Genome Operation

Lecture Week 8:

Genetic code expansion (GCE):

POI-site-specific GCE and whole-proteome-based GCE tools

2023 Nov 7th (Room: BS 270): 10:15 am – noon

https://moodle.epfl.ch/course/view.php?id=15521

Part I: Chem Bio Toolsets

Chemical biology tools

Week	Date	Topic	Notes
1	19th Sept	Intro + Enzyme inhibitors as drugs	
2	26th Sept	Tools modulating enzyme activity/protein function	PSet 1
3	3rd Oct	Genetic vs. pharmacological perturbation methods	
4	10th Oct	Quantitative mass spectrometry: concepts & tools	PSet 2
5	17th Oct	Leveraging quant. Proteomics for target profiling	
6	24th Oct	Tools interrogating the genome & transcriptome	PSet 3
7	31st Oct	… topic above continues	
8	7th Nov	Tools interrogating proteome-level processes	PSet 4

Biotechnological Applications using tools hijacking *nascent* protein (poly-peptide) translation

Continued from the last lecture….

- How do we **rapidly evolve/engineer proteins/enzymes** and target-ligand pairs?

mRNA display technology

- How can we site-specifically integrate an unnatural amino acid **(expanding the genetic code)** such as to **understand/track the precise impact of a single-amino acid switch on a single protein of interest?**

POI-site-specific GCE

- How do we **monitor time-resolved changes in global protein synthesis/turnover**?

Whole-proteome GCE

Genetic code expansion (GCE) technology
(aka: amber suppression method; unnatural amino acid (UAA) incorporation technology)

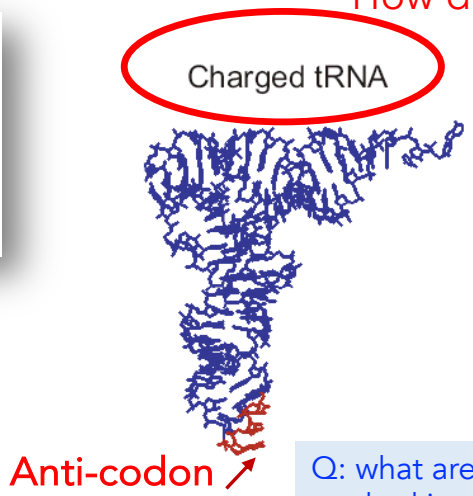
POI: protein of interest

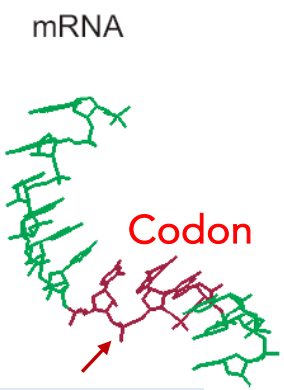
aka: also known as

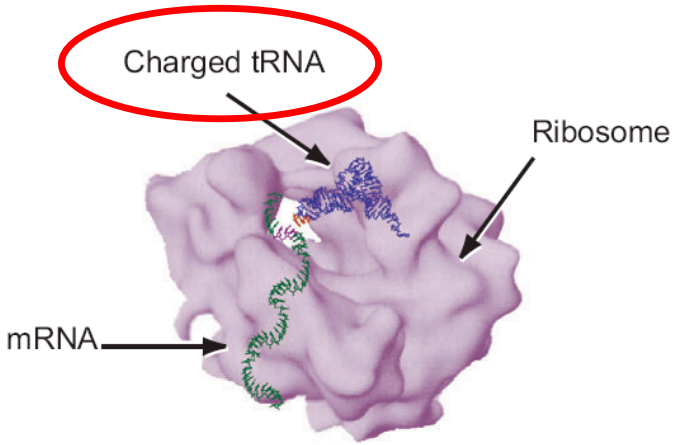
Amber suppression [aka, Genetic Code Expansion (GCE)] technology: Hijacking nascent mRNA translation

DNA
⇩
mRNA (codons)
⇩
tRNA
(anti-codon loop)
⇩
PROTEIN

How do we get to 'charged tRNA'?



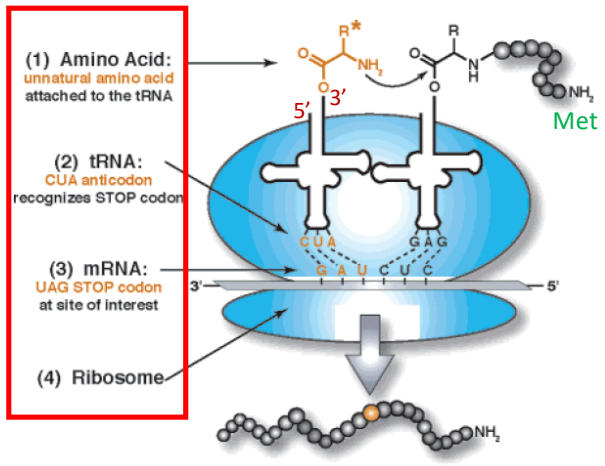




Q: what are these 2 'regions' marked in red?

Termination codons (amber, ochre, opal) IMPLICATION of ONLY needing a single STOP codon:

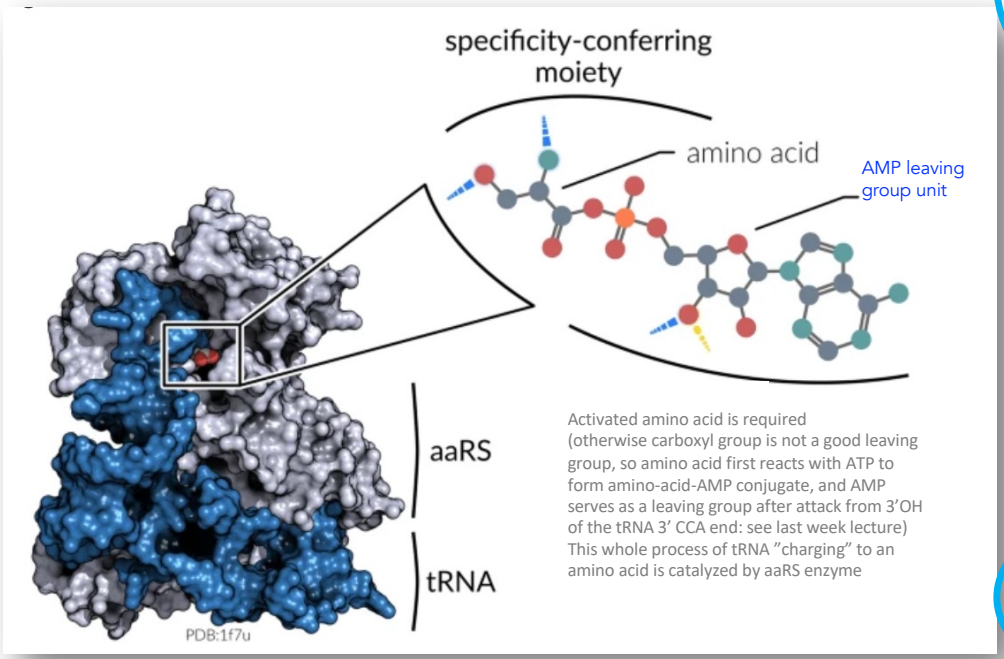
Suppressor tRNA can be "charged" with an unnatural amino acid



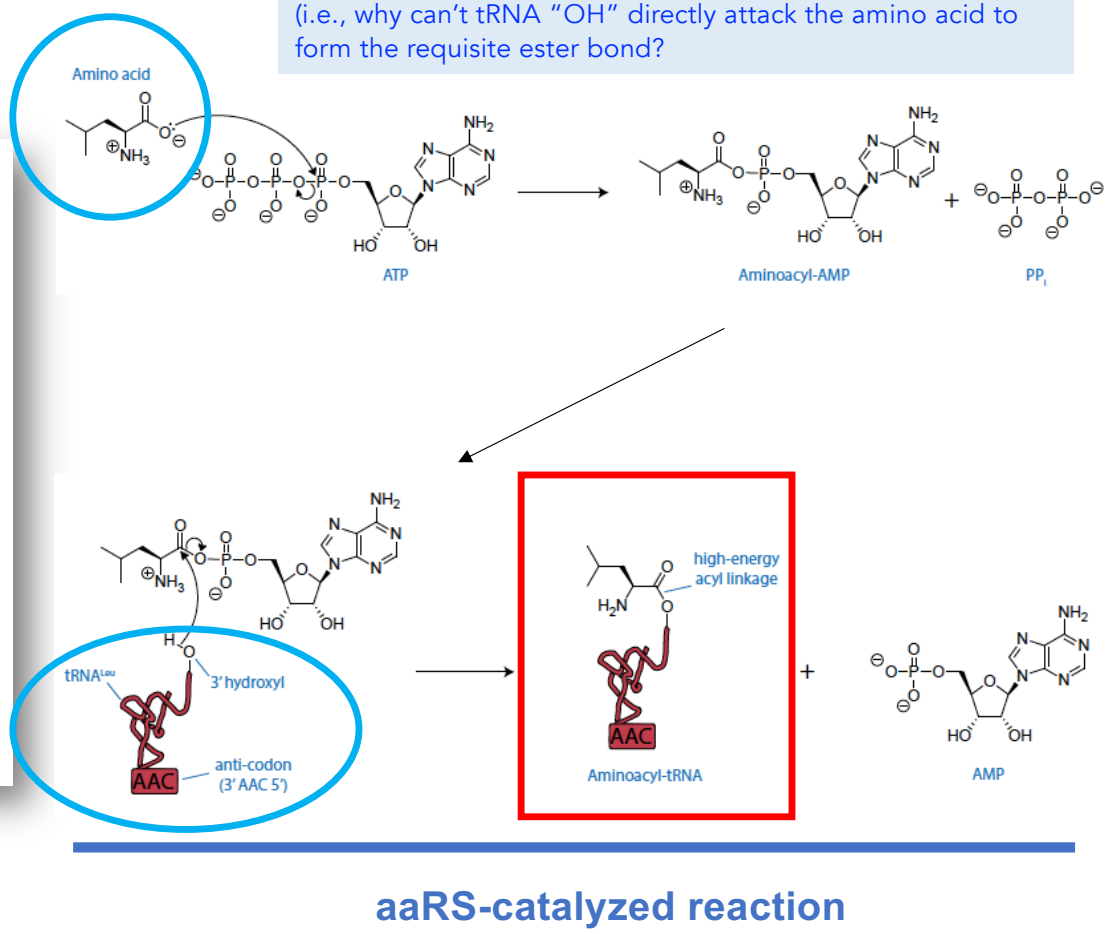
Slide from last lecture

The chemistry behind tRNA charging process

3D structure of a representative aaRS·tRNA complex



Q: why is ATP-activation process of the amino acid necessary? (i.e., why can't tRNA "OH" directly attack the amino acid to form the requisite ester bond?



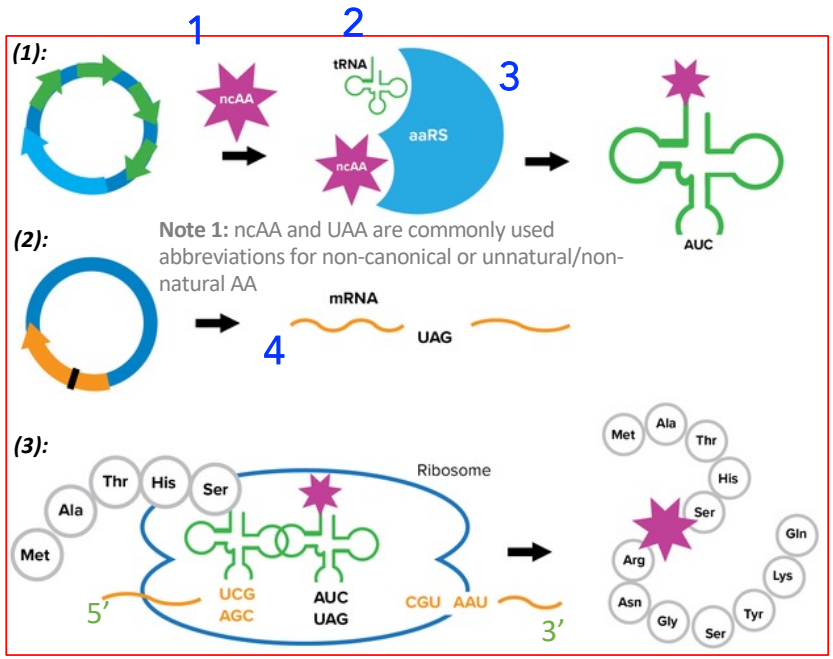
aaRS-catalyzed reaction

aaRS = aminoacyl tRNA synthetase

Slide from last lecture

Genetic code expansion: the required components

Adapted from: https://www.addgene.org/collections/genetic-code-expansion/



To expand the genetic code, modified tRNAs, codons, and tRNA synthetases are introduced into the cell on plasmids and the new amino acid is introduced in the media. We generally need two plasmids (see 1 and 2 in figure):

__1:__ A plasmid expressing the tRNA and its cognate aminoacyl-tRNA-synthetase (aaRS) that has been evolved to incorporate non-canonical amino acids (ncAAs).

__2:__ A plasmid encoding your favorite gene with the modified codon (typically the amber codon) (integrated at the right site to ultimately integrate the ncAA) that is to be recognized by the cognate charged tRNA.

__3:__ Once these plasmids have been introduced in the cells, the non-canonical amino acid can be incorporated using the existing protein translation machinery.

To expand the genetic code, 4 major changes to the standard translation machinery are needed in order to incorporate a non-canonical amino acid into the protein of interest:

--The non-canonical amino acid, which is generally introduced in the media.

-- A new codon to be allocated to the new amino acid. Because there are no free codons, this can be challenging. In *E.coli*, the rarest codon is the amber stop codon (UAG) and thus this codon is often used. The gene of interest can be expressed from a plasmid containing a UAG codon at the place where the new amino acid would be incorporated. Other options, such a 4-base pair codons, have also been utilized.

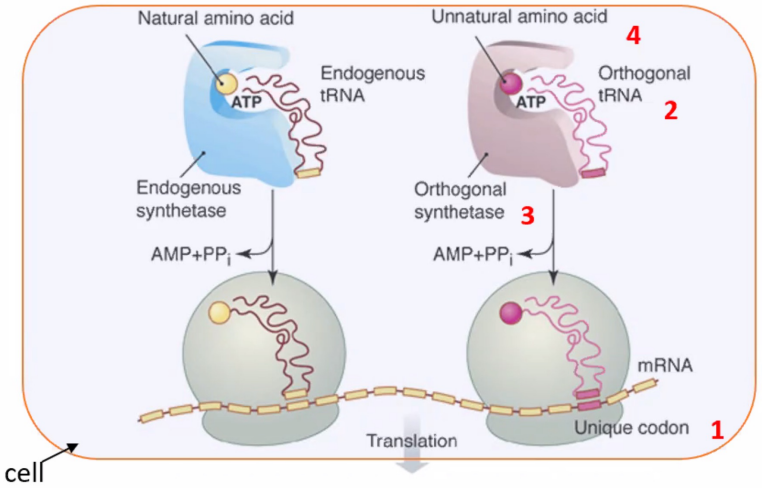
-- A tRNA that recognizes this codon.

-- An aminoacyl-tRNA synthetase (aaRS) to load the new amino acid onto the tRNA. The tRNA and synthetase are called an orthogonal set, because they should not crosstalk with the endogenous tRNA and synthetase sets. Many of these sets are derived from *M. jannaschii*, *M. barkeri*, or *E.coli* and can be mutated and screened through directed evolution to charge the tRNA with a different amino acid. They are typically expressed from a single plasmid, with multiple copies of the tRNA.

Unnatural Amino Acids (UAAs) incorporated via GCR (or amber suppression) method

INTERLUDE

Reprogramming the Translational Machinery: A Molecular Problem



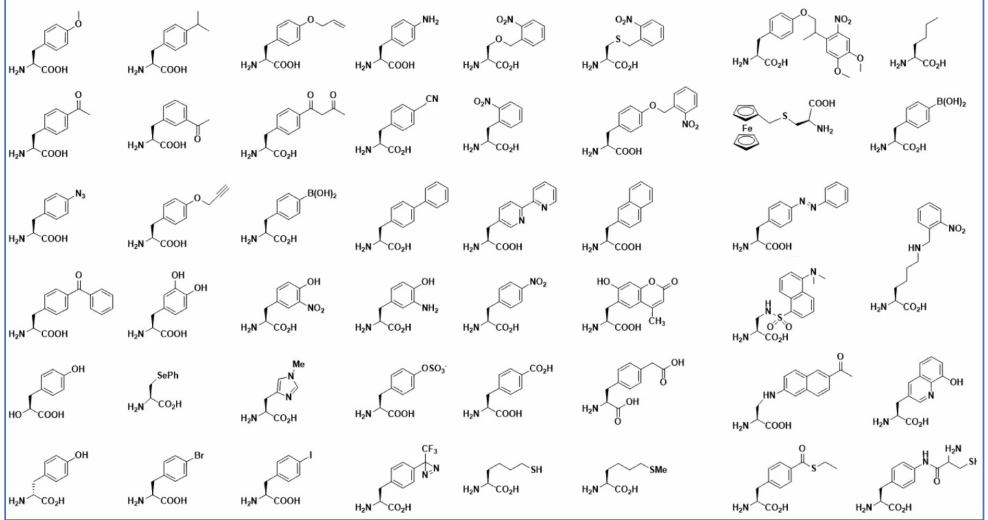
New Components

(1) "Blank" codons (nonsense codon, four base codon, repurposed codons, etc.)
(2) Engineer a cognate tRNA that is orthogonal to all other aminoacyl-tRNA synthetases
(3) Evolve an orthogonal aminoacyl-tRNA synthetase to uniquely recognize this tRNA
(4) Evolve this synthetase to uniquely charge this tRNA with the 21st amino acid
(5) Biosynthesize or transport the amino acid into the cell

➢ Roughly 200 unnatural amino acids with high translational fidelity and good yields (up to 8 grs/L and > 2,000 L scale) in bacteria, yeast and mammalian cells Young, Schultz ACS Chem Bio 2018

The Expanding Genetic Code

Example UAAs successfully incorporated:



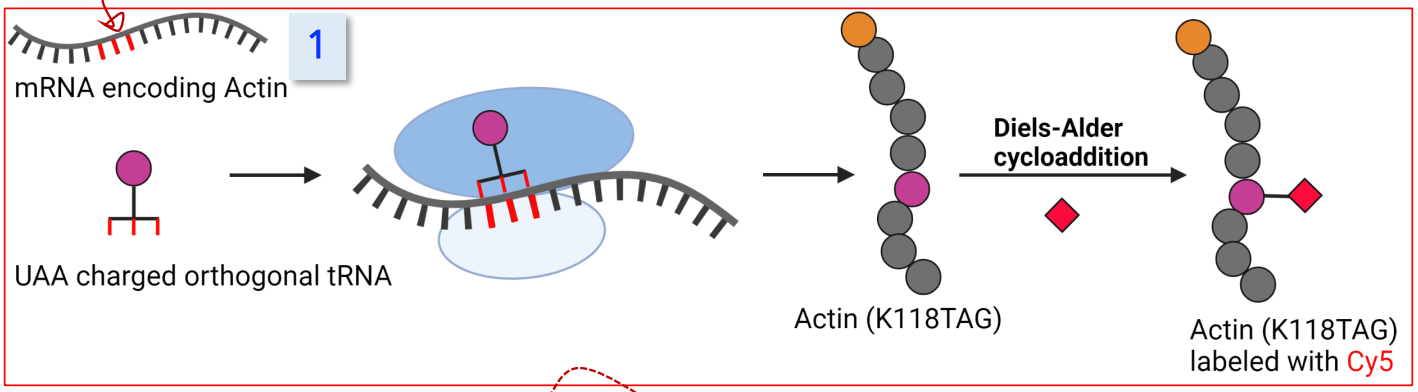
Relevant slides presented by Prof. Peter Schultz (pioneer of the GCE technology) during 2021 Intn'l Chem Bio Society conference, a leading pioneer in GCE method development and applications.

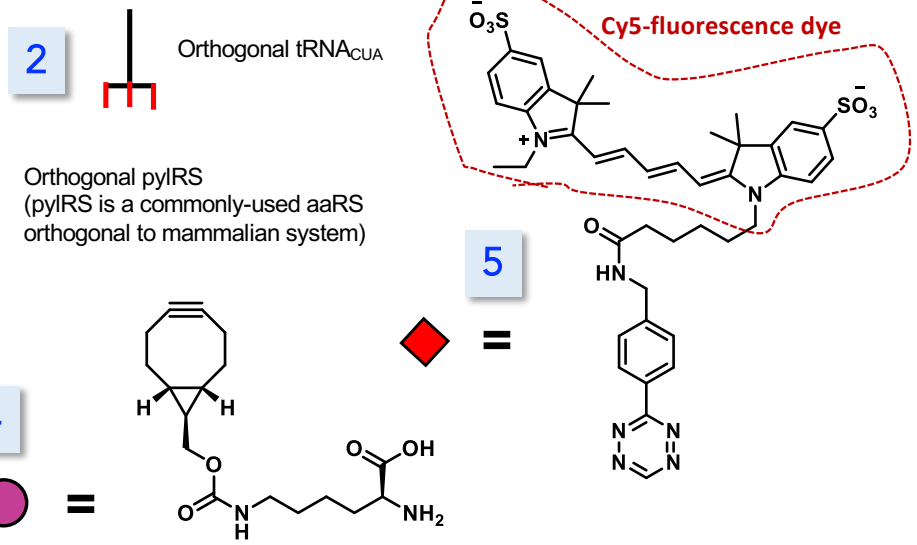
In-class Exercise: Fill out the missing fluorescence gel and western blot data

Full-length actin protein is made up of 374 amino-acids and its molecular weight (MW) is 42 kDa

HA Actin (K118TAG) ← note: TAG at DNA level

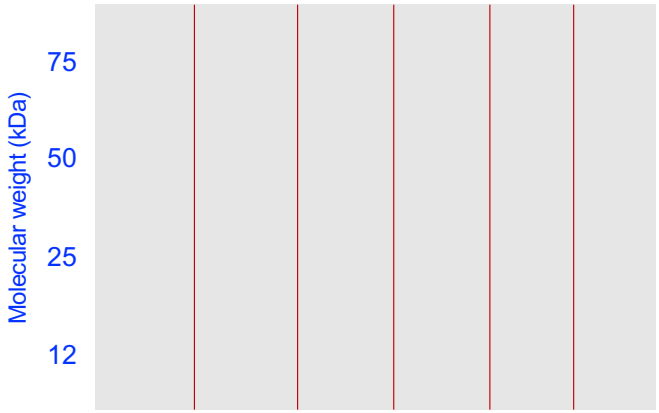
UAG at mRNA level



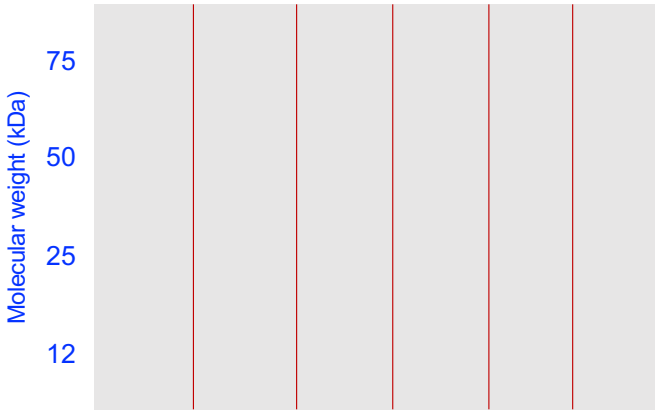


Experimental setup:
Cells were subjected to different conditions shown, such that they express 1, 2, and/or 3, and either treated or untreated with compound 4 for 5 hrs. Cells were then lysed and compound 5 was added to the lysate. The samples were then analyzed by SDS-PAGE and imaged detecting either Cy5 signal (top) or anti-HA signal (bottom)

1.	+	+	+	+	-	+
2.	+	+	+	-	+	+
3.	+	+	-	+	+	+
4.	+	-	+	+	+	+
5.	-	+	+	+	+	+



Fluorescence gel (Cy5 signal)



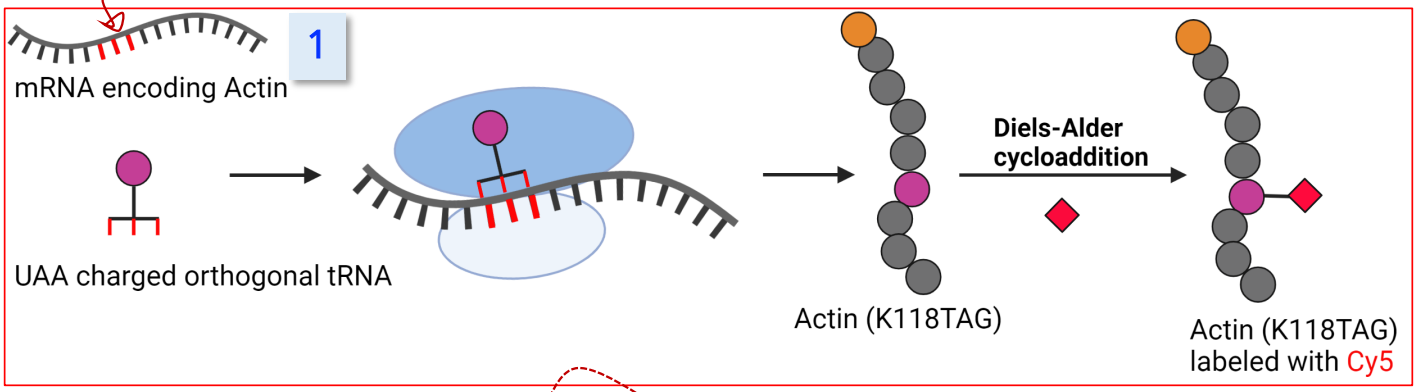
Western blot (anti-HA signal)

In-class Exercise: Fill out the missing fluorescence gel and western blot.

Full-length actin is made up of 374 amino-acids and its molecular weight (MW) is 42 kDa

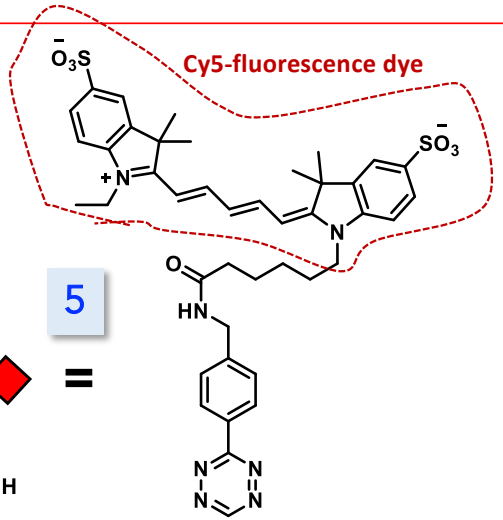
HA Actin (K118TAG) ← note: TAG at DNA level

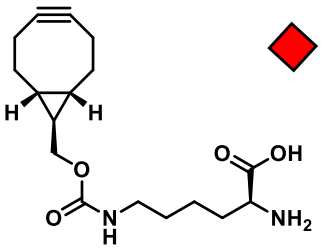
UAG at mRNA level



2 Orthogonal $tRNA_{CUA}$

3 Orthogonal pylRS (pylRS is a commonly-used aaRS orthogonal to mammalian system)

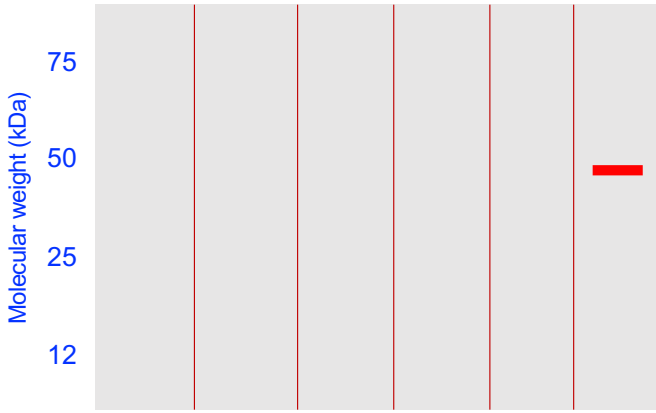


4 = 

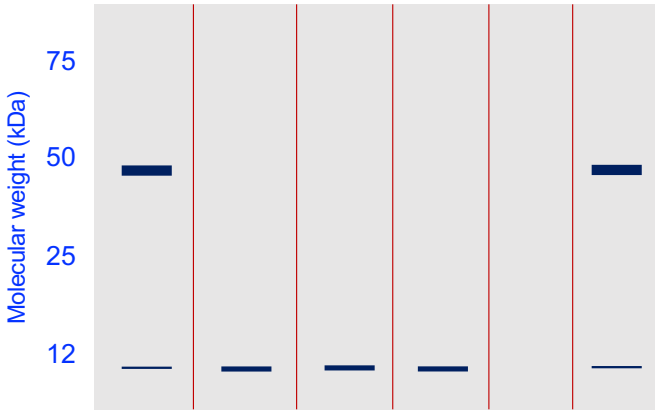
Experimental setup: Cells were subjected to different conditions shown, such that they express 1, 2, and/or 3, and either treated or untreated with compound 4 for 5 hrs. Cells were then lysed and compound 5 was added to the lysate. The samples were then analyzed by SDS-PAGE and imaged detecting either Cy5 signal (top) or anti-HA signal (bottom)

KEY

1.	+	+	+	+	-	+
2.	+	+	+	-	+	+
3.	+	+	-	+	+	+
4.	+	-	+	+	+	+
5.	-	+	+	+	+	+



Fluorescence gel (Cy5 signal)



Western blot (anti-HA signal)

Biotechnological Applications using tools hijacking *nascent* protein (poly-peptide) translation

- How do we **rapidly evolve/engineer proteins/enzymes** and target-ligand pairs?

mRNA display technology

- How can we site-specifically integrate an unnatural amino acid **(expanding the genetic code)** such as to **understand/track the precise impact of a single-amino acid switch on a single protein of interest?**

POI-site-specific GCE

- How do we **monitor time-resolved changes in global protein synthesis/turnover**?

Whole-proteome GCE

Genetic code expansion (GCE) technology
(aka: amber suppression method; unnatural amino acid (UAA) incorporation technology)

POI: protein of interest

aka: also known as

The BONCAT strategy for labeling, detection and identification of newly synthesized proteins.

Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging

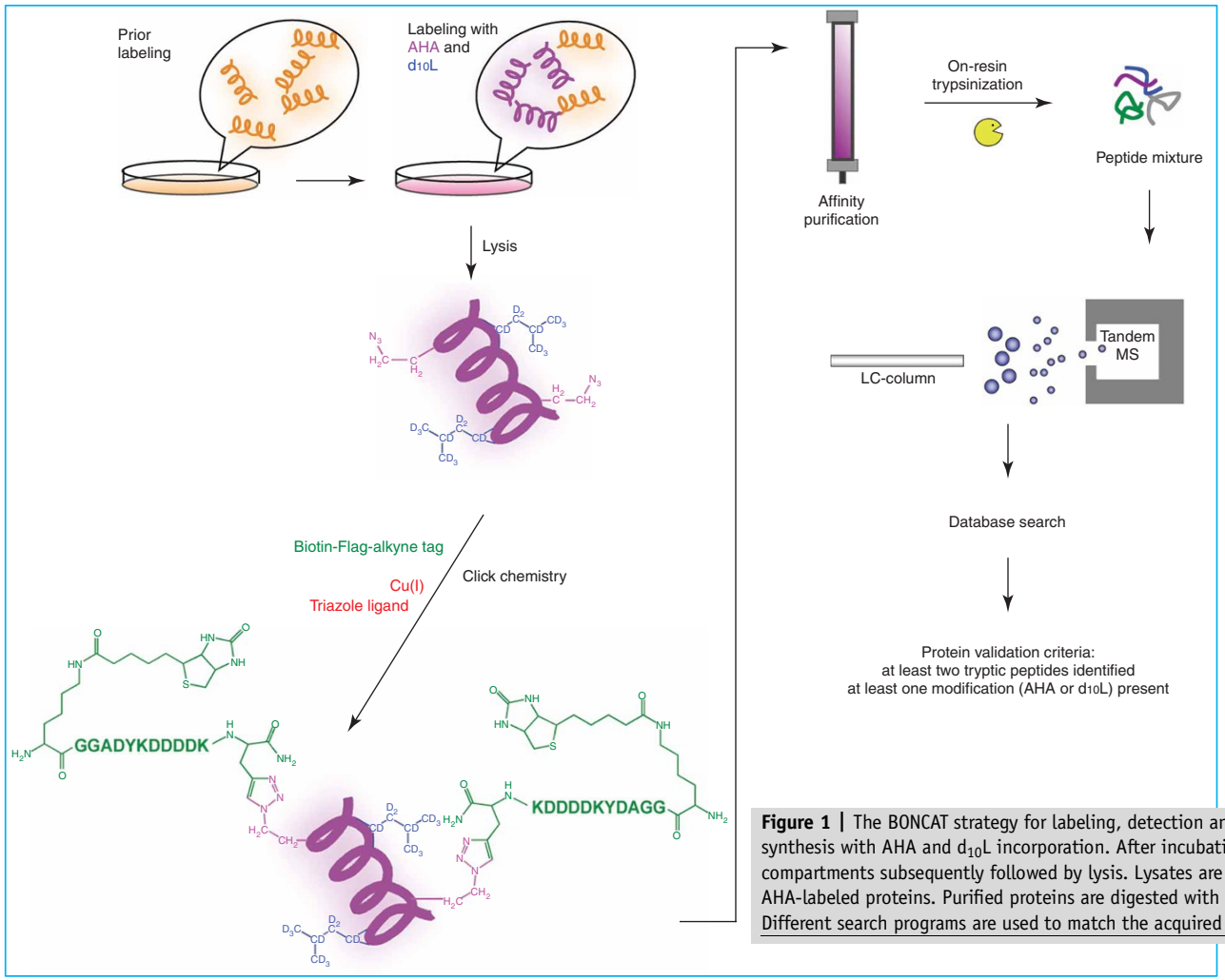
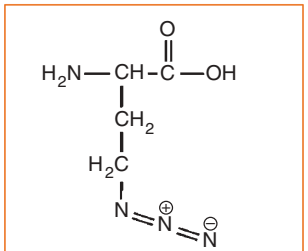


Figure 1 | The BONCAT strategy for labeling, detection and identification of newly synthesized proteins. Cells are incubated with AHA and d_{10}L to allow protein synthesis with AHA and d_{10}L incorporation. After incubation, cells are lysed or undergo a subcellular fractionation for biochemical enrichment of specific cellular compartments subsequently followed by lysis. Lysates are then coupled to an alkyne-bearing affinity tag, followed by affinity chromatography, to enrich for AHA-labeled proteins. Purified proteins are digested with a protease and the resulting peptides are analyzed by tandem MS to obtain experimental spectra. Different search programs are used to match the acquired spectra to protein sequences.

Q: could we simply replace a given natural amino acid fed to the cell growth media with an unnatural mimic?

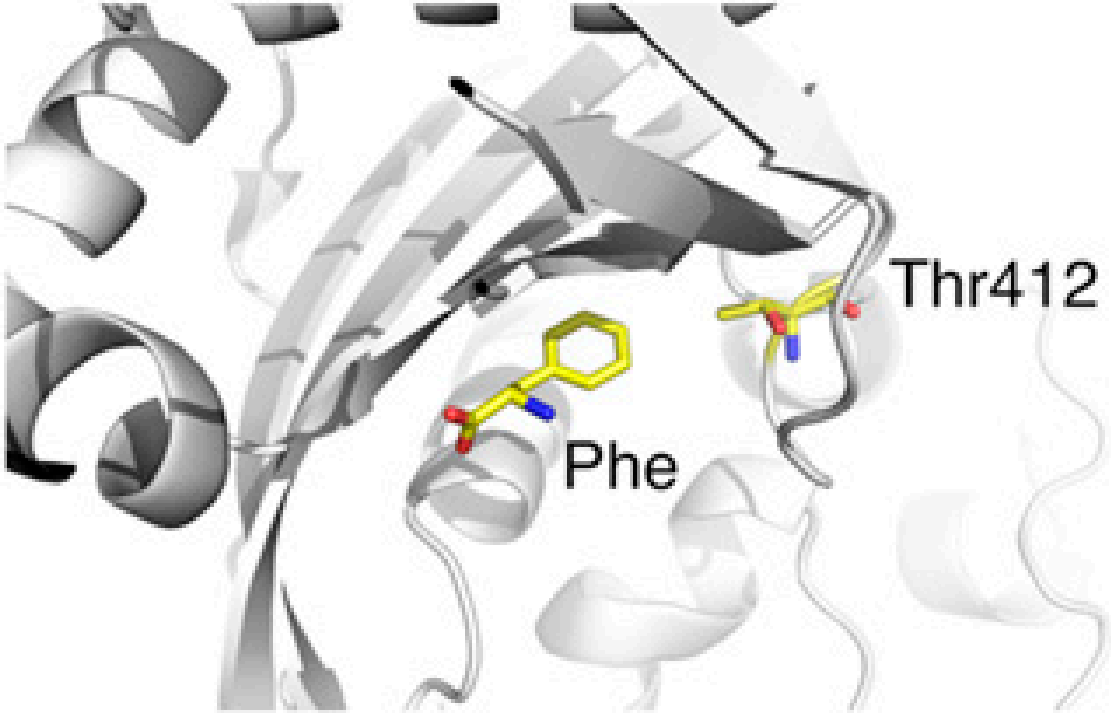


Here, deuterated leucine (d_{10}L) -based metabolic co-labeling approach was adopted such that candidate proteins can be simply validated by tandem MS

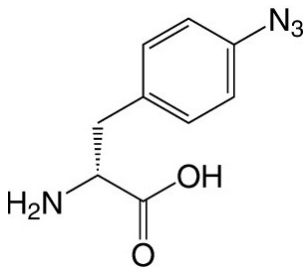
Q: could we achieve a more efficient whole-proteome labeling by an UAA by introducing an orthogonal aaRS specifically tailored for that UAA?

General design strategy for orthogonal aaRSs that can accommodate UAA

PNAS | **March 3, 2015** | vol. 112 | no. 9 | **2705–2710**



Azido-Phe:

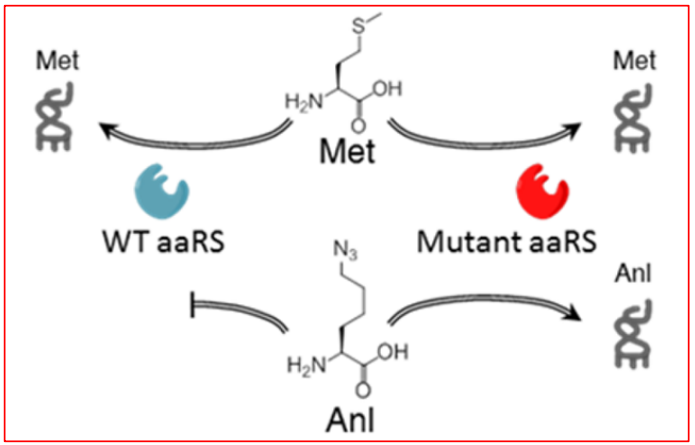


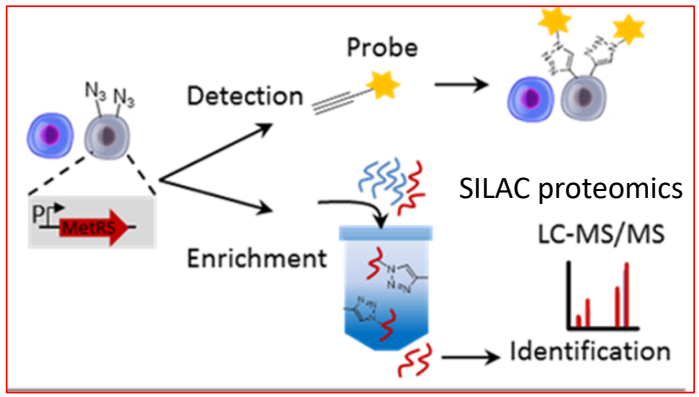
Q: what is the rational in mutating Thr412 to Gly (a smallest natural AA residue available)?

Active site of eukaryotic PheRS with bound Phe is shown above

An evolved aaRS that can efficiently catalyze formation of appropriate orthogonal tRNA charged with Azido-Phe was successfully achieved by mutating Thr412 to Gly

SILAC – BONCAT technique & applications : In-Class Exercise

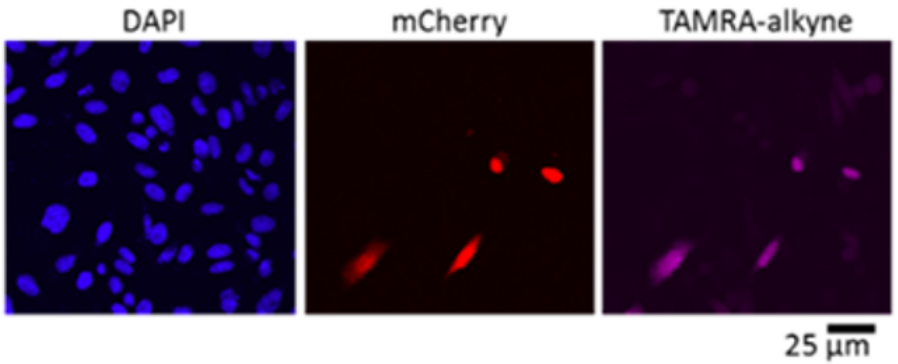


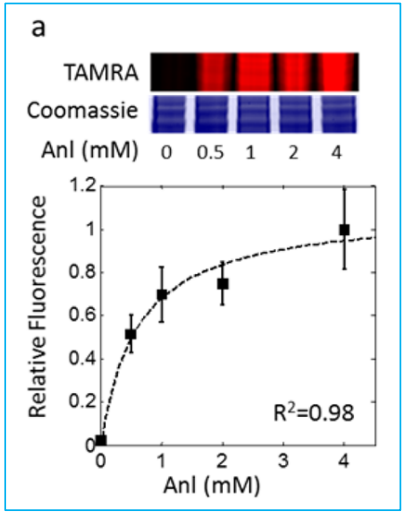


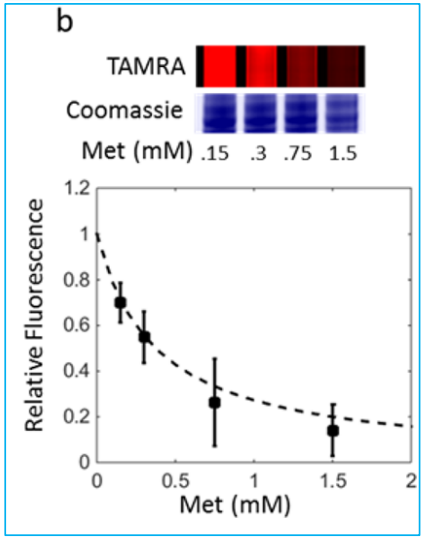
c

Selectivity against endogenous canonical (wt) aaRS is determined (see Figure a and b):
a) shows whole-proteome labeling in lysates derived from cultured cells expressing the mutant MetRS treated with different Anl concentrations shown.
b) shows whole-proteome labeling in above cells treated with increasing dosage of Met but fixed dose of Anl
Evaluate the two data sets and comment on the selectivity.

In figure c: Live cells containing the shown construct were imaged. Note: T2A is a self-cleaving peptide that allows expression of two separate proteins; here, MetRS and mCherry. TAMRA is a synthetic small-molecule fluorescence dye (here conjugated to alkyne). DAPI is the stain for nuclei. Comment on the data.







DOI: 10.1021/jacs.5b08980
J. Am. Chem. Soc. 2016, 138, 4278–4281

Learning outcomes (Week 8: CH-313 Chemical Biology - Synopsis)

- Hijacking protein translation in Genetic Code Expansion (GCE) technology
- Applications of GCE, including experimental design, data interpretations, as well as benefits and limitations of GCE (both POI-site-specific GCE and whole-proteome-based GCE, often referred to as BONCAT method)