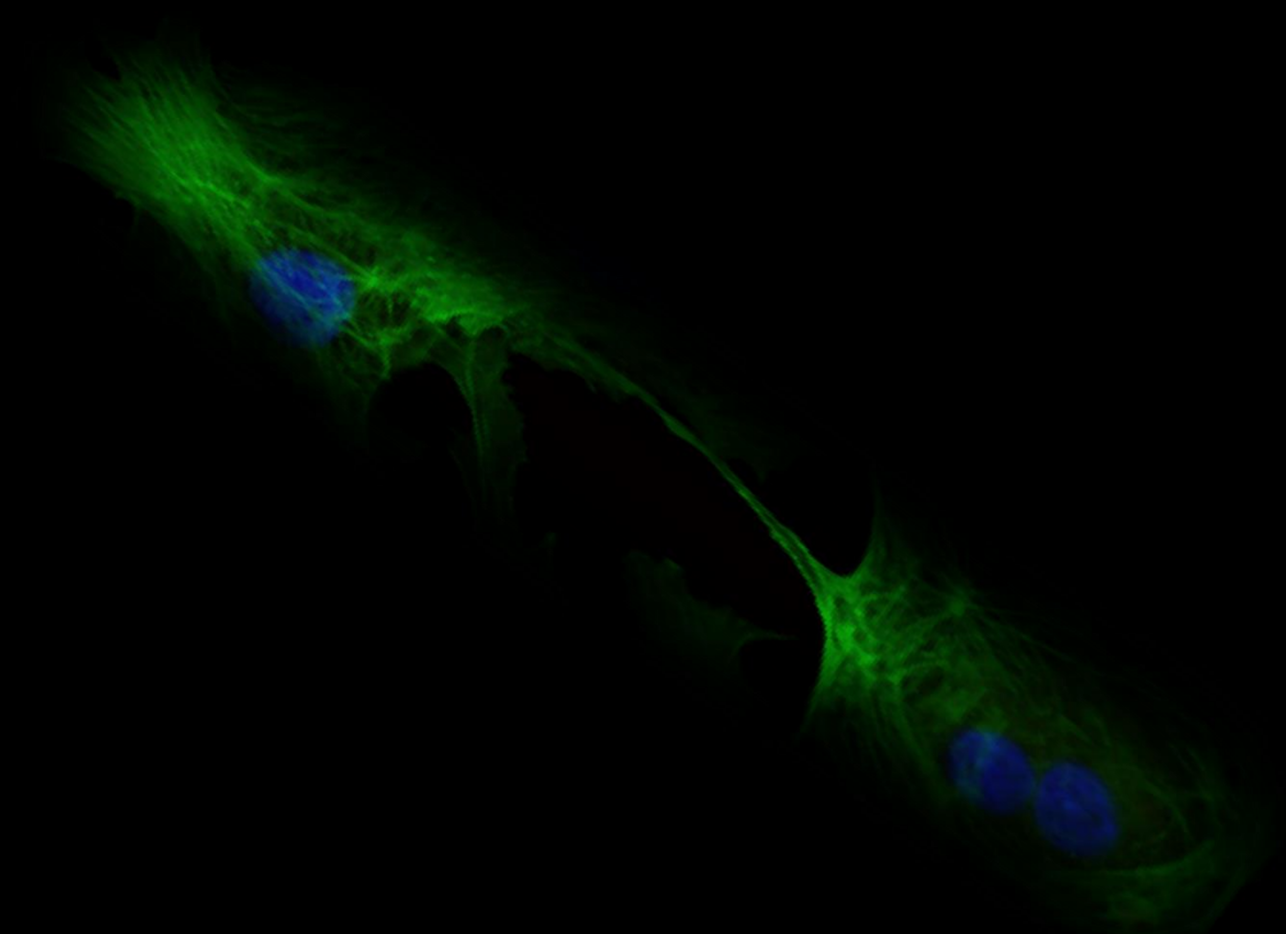
# 3. Implementation of chemical reporters



# Bioorthogonal modules

- *Target biomolecule* : protein, lipid, nucleic acid etc.
- *Chemical reporter* : non-native, non-perturbing chemical handles that can be modified in living systems through highly selective reactions with exogeneously delivered probes
- *Signaling unit* : fluorescent, radioactive, NMR active nuclides, PET etc.

J. A. Prescher, C. R. Bertozzi (2005). *Nat. Chem. Biol. 1*, 13

# Bioorthogonal schemes

- Two-step labeling strategy
- Incorporation of a bioorthogonal function
- Modulation with a signaling unit

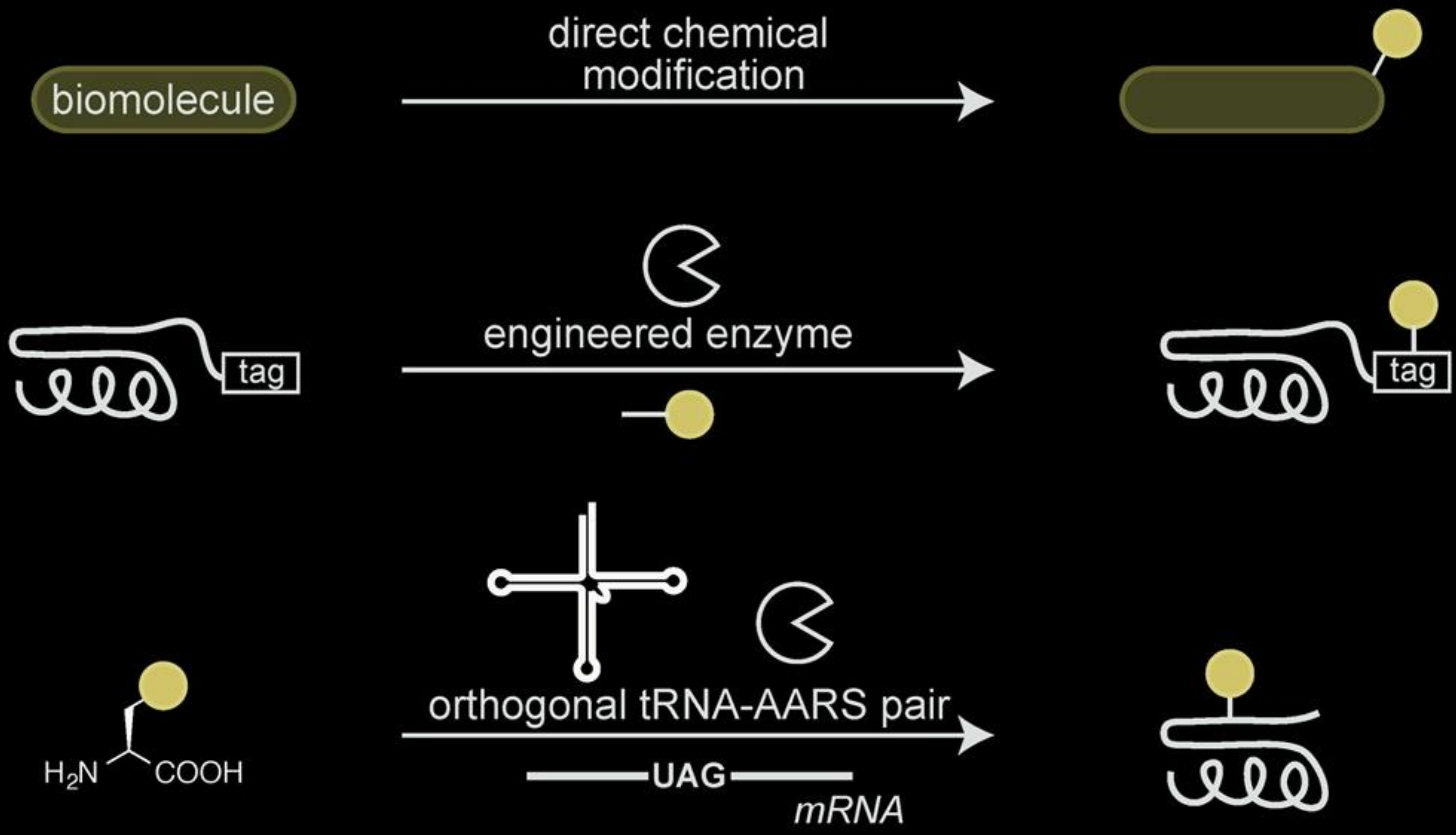
J. A. Prescher, C. R. Bertozzi (2005). *Nat. Chem. Biol. 1*, 13

# Implementation of chemical reporters

- Metabolism (bioorthogonalized metabolites)
- Genetic encoding (non-canonical amino acids)
- Enzymatically with engineered enzymes (e.g. PCR)
- Chemical ligation (unique natural motifs)


- C. R. Becer, R. Hoogenboom, U. S. Schubert (2009) *Angew. Chem. Int. Ed*. *48*, 4900
- K. E. Beatty (2011) *Mol. BioSyst. 7*, 2360

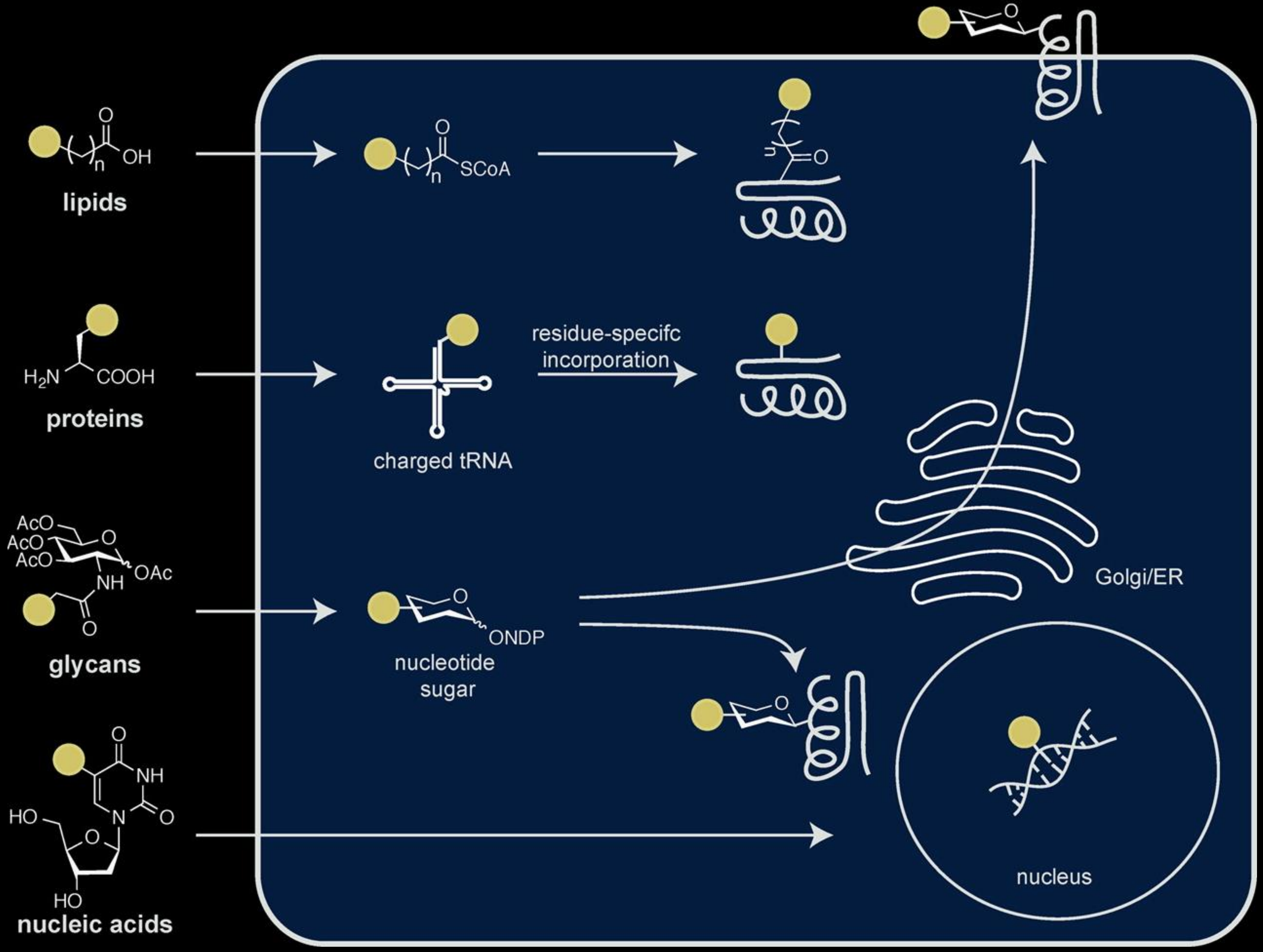
# Installation methods




D. M. Patterson et al. (2014) *ACS Chem. Biol. 9*, 592

# Installation methods






D. M. Patterson et al. (2014) *ACS Chem. Biol. 9*, 592

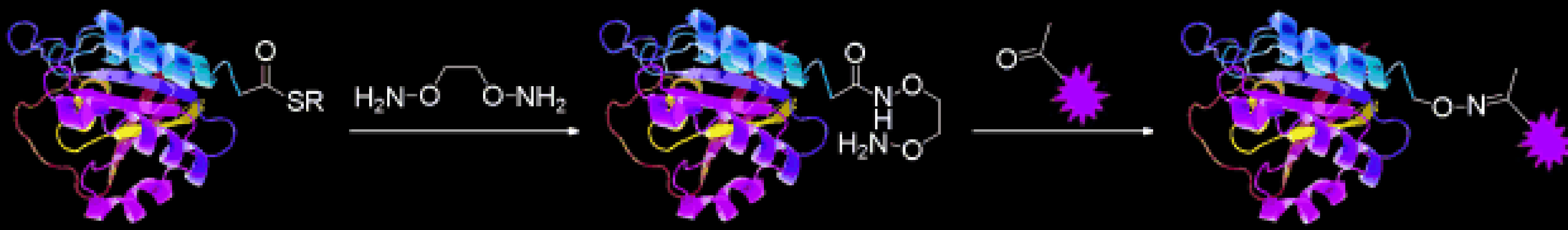
# 3.1 Direct chemical modification

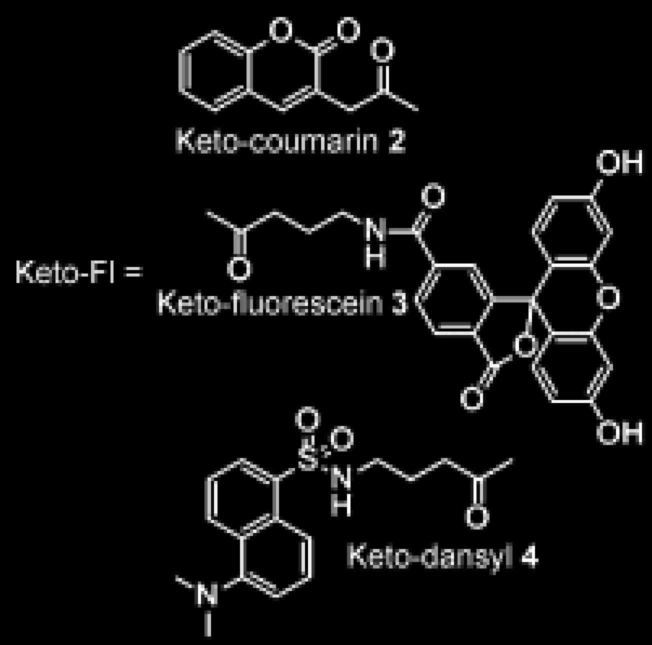
- Efficient, fast
- Non-selective, non site-specific
- Mostly proteins
- N- or C-termini (different pKa)
- Cys, Lys, aromatic side chains

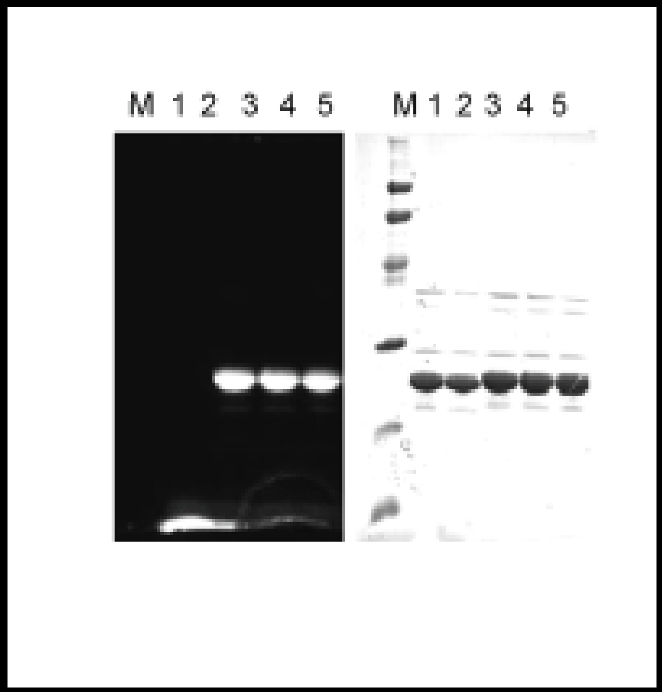
D. M. Patterson et al. (2014) *ACS Chem. Biol. 9*, 592

# C-terminal modification (with bis-oxyamine)





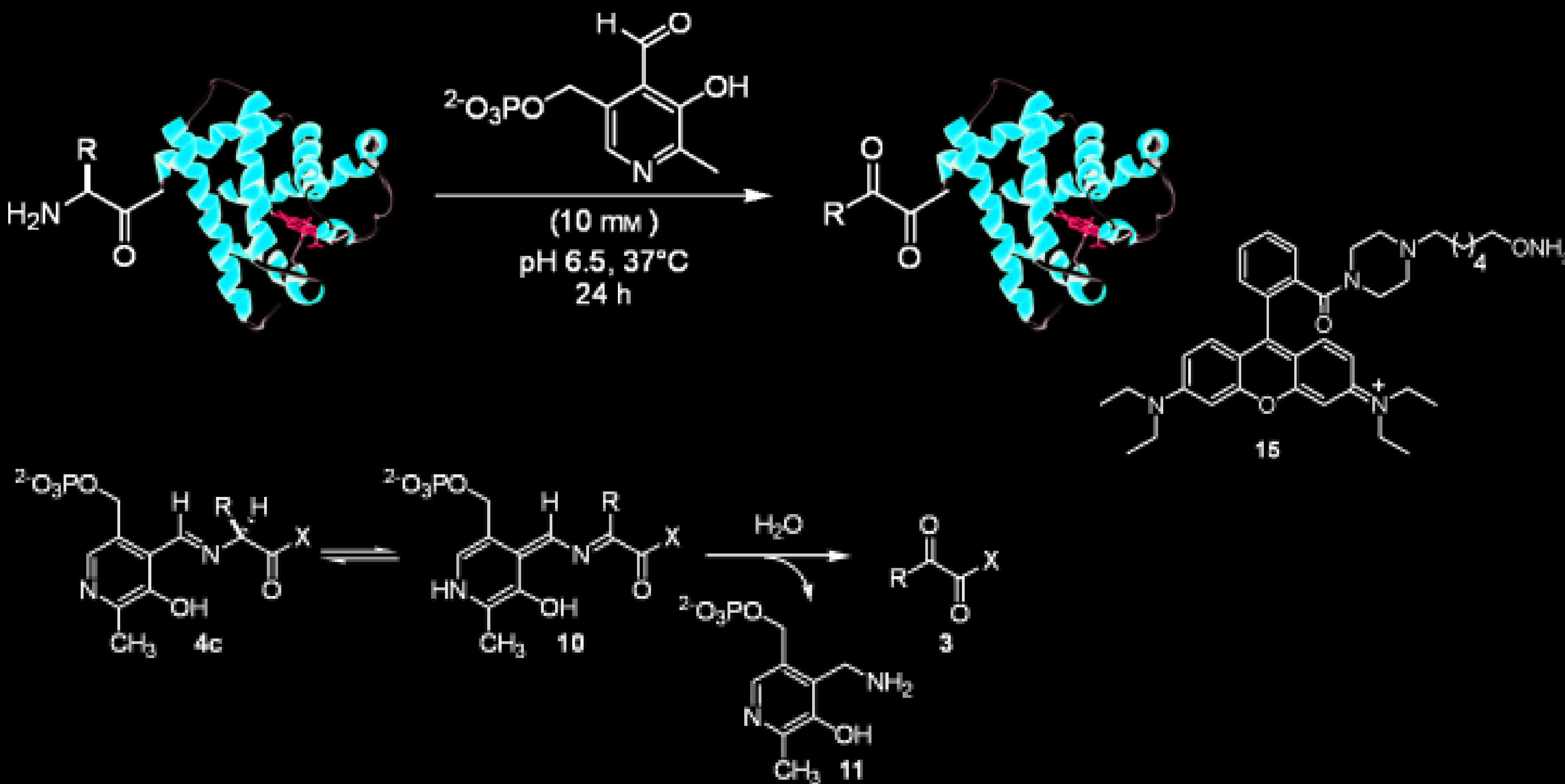




L. Yi et al. (2010) *Angew. Chem. Int. Ed. 49*, 9417

# N-terminal modification (with pyridoxal-phosphate)



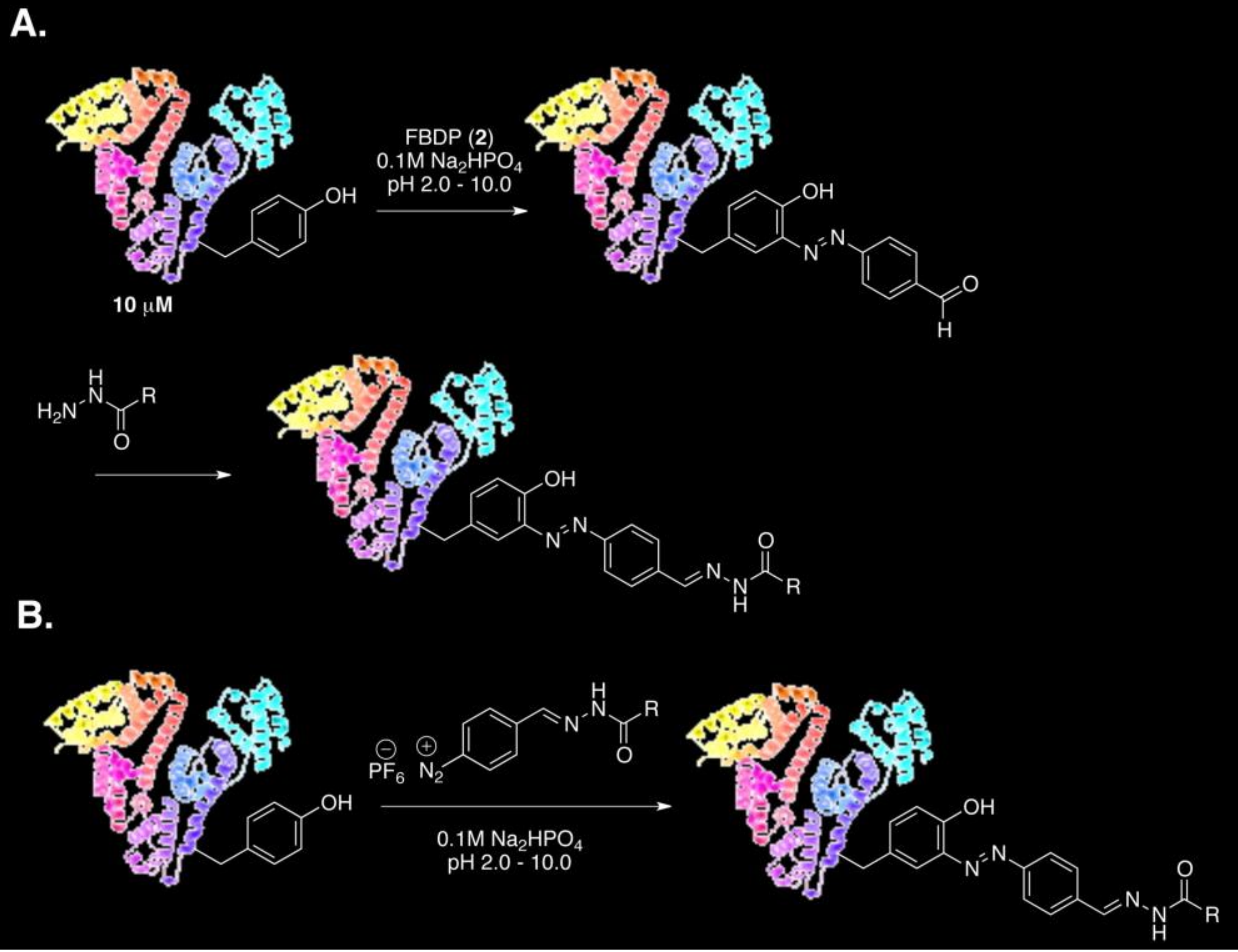



- Compatible with thiols

J. M. Gilmore et al. (2006) *Angew. Chem. Int. Ed. 45*, 5307

# Tyrosine modification

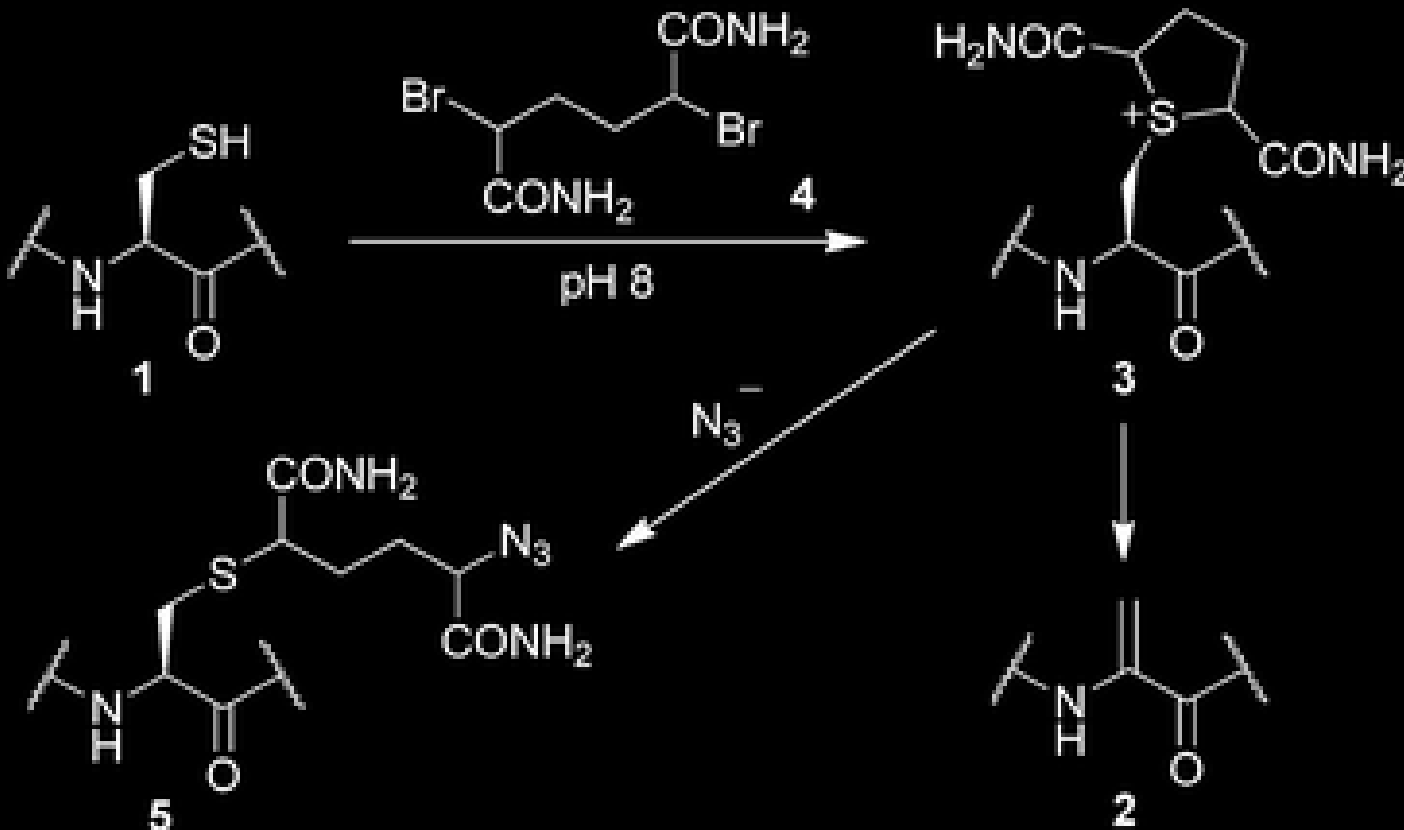





J. Gavrilyuk et al. (2012). *Bioconjugate Chem.*, *23*, 2321

# Cysteine modification with dibromohexanediamide

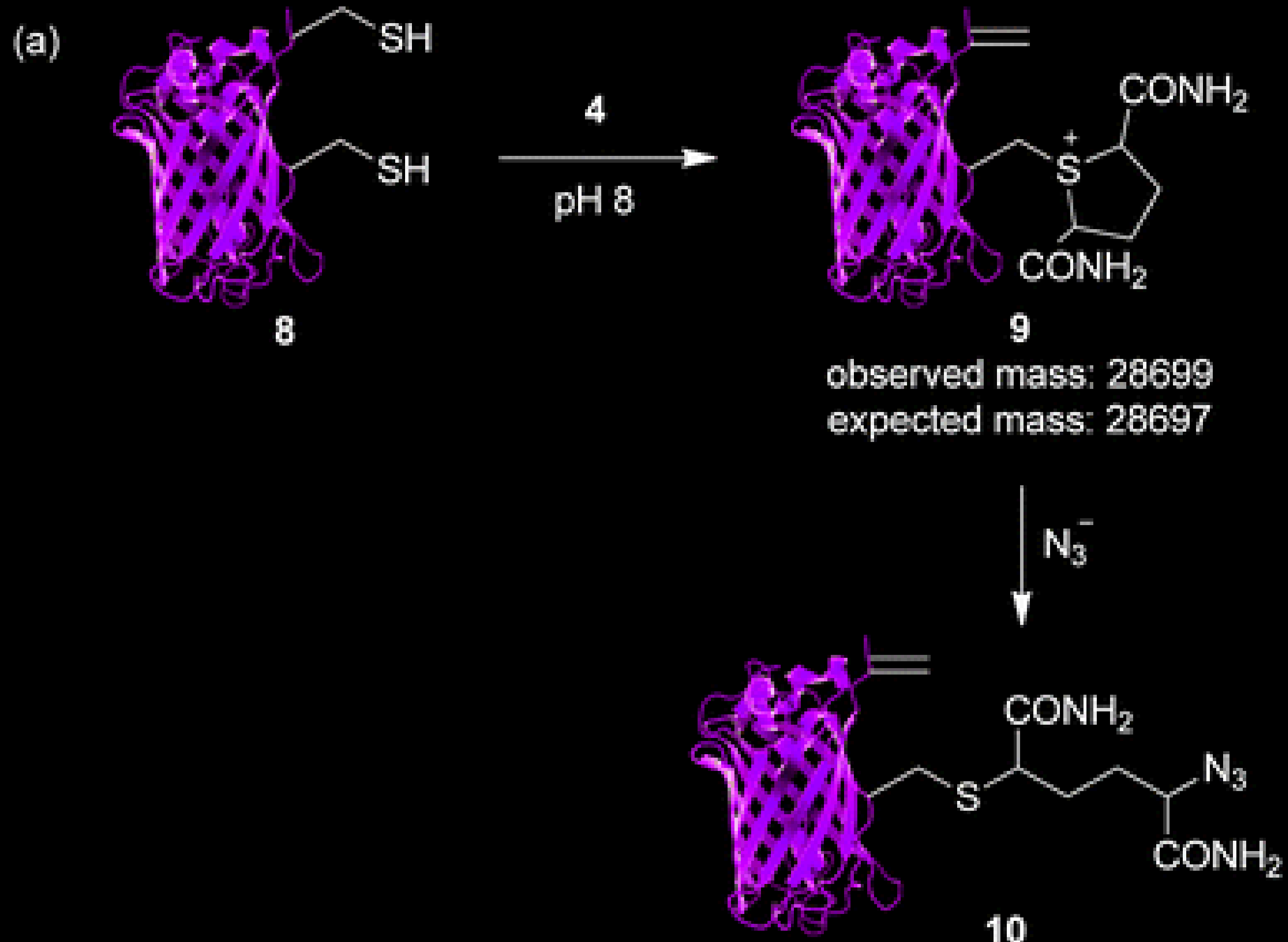





Environment dependent reactions

R. I. Nathani et al. (2013). *Chem. Sci.*, *4*, 3455

# Cysteine modification




R. I. Nathani et al. (2013). *Chem. Sci.*, *4*, 3455

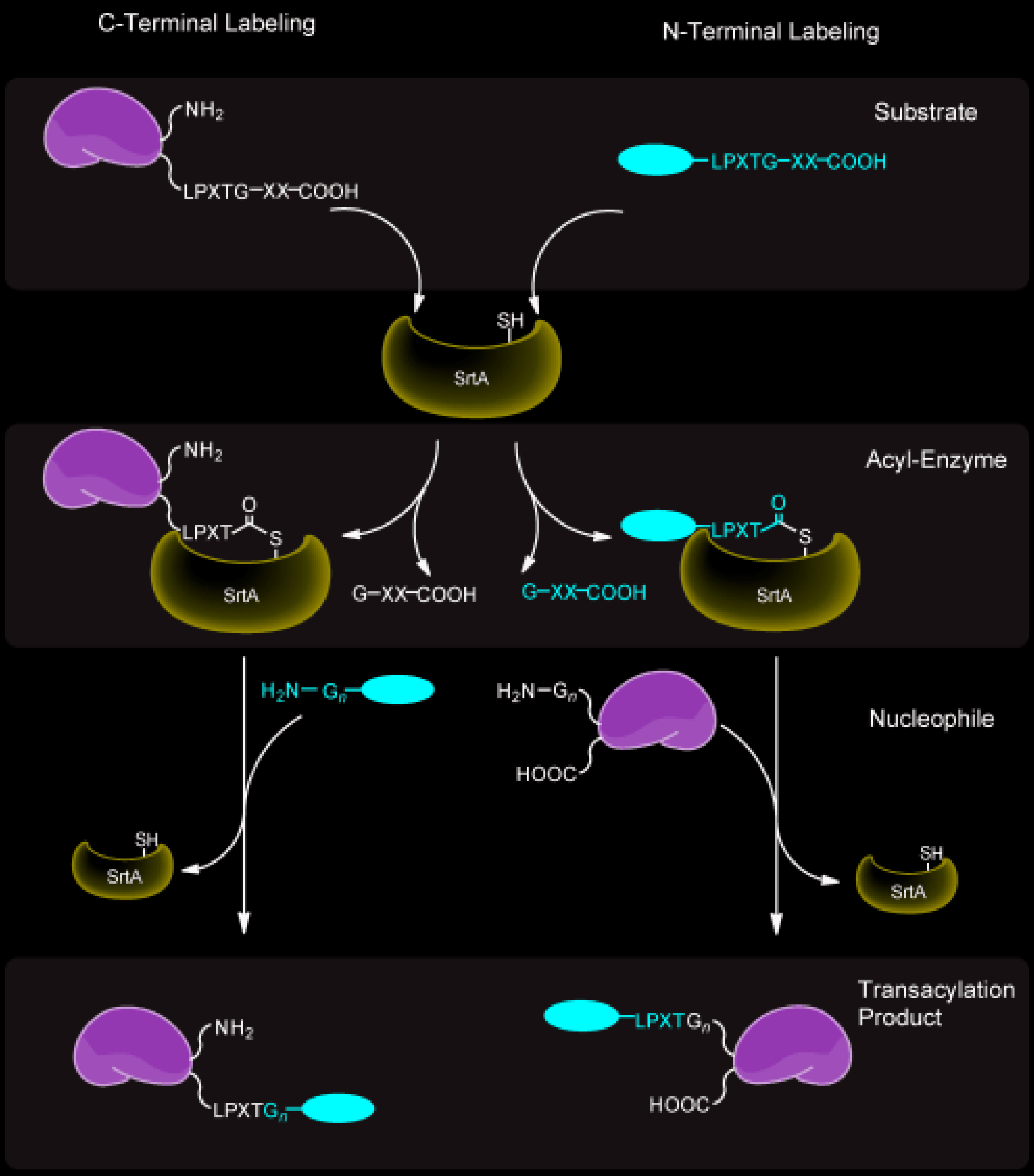
# 3.2. Enzymatic tagging

- Enzymatic platforms
- Mostly ligases, transpeptidases
- Modified substrates
- Recognizing consensus sequences

D. M. Patterson et al. (2014) *ACS Chem. Biol. 9*, 592

# Sortase based tagging





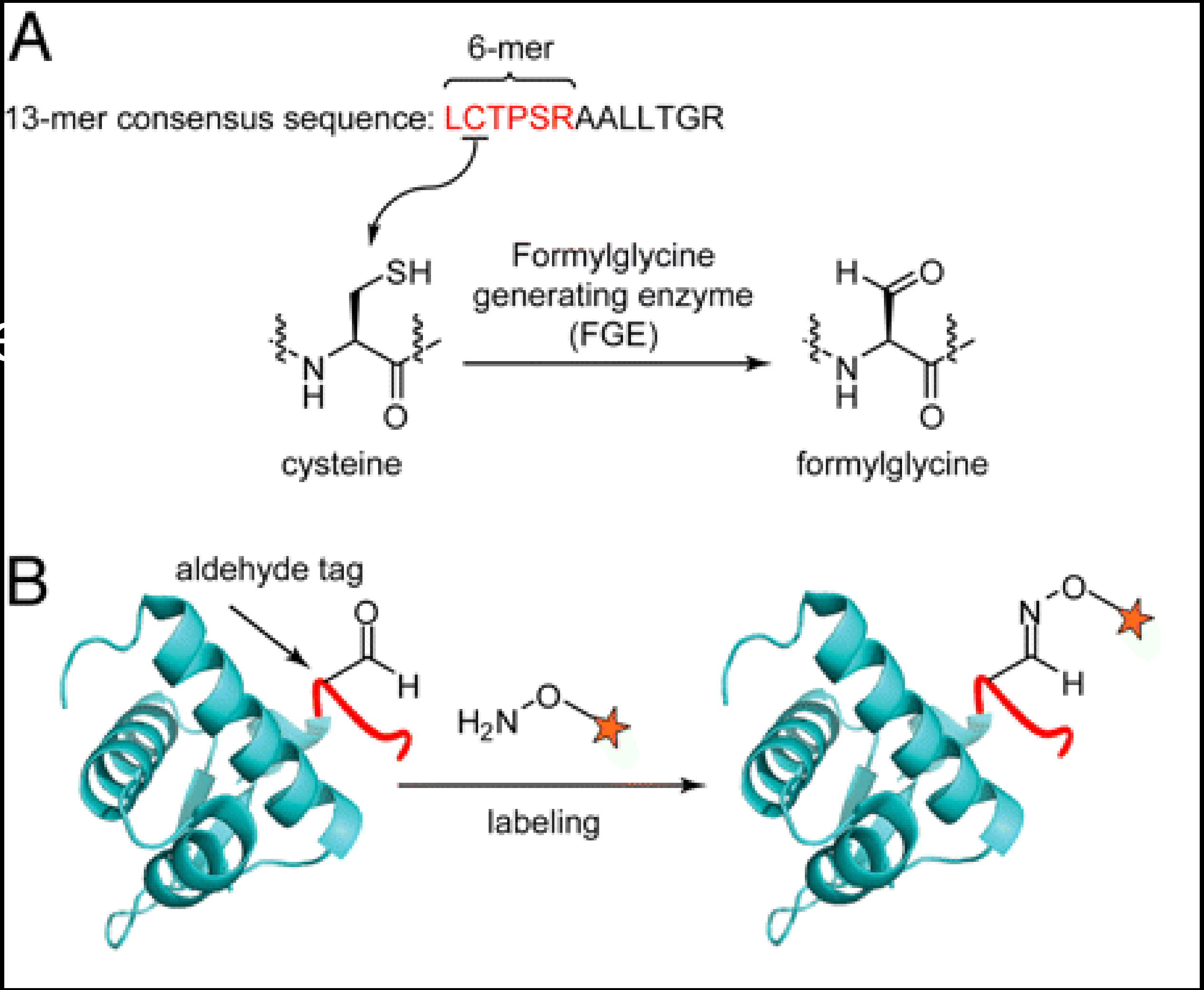

- Site specific bond-breaking
- Reform a bond
- Thiol-transpeptidase

M. W. Popp et al. (2011). *Angew. Chem., Int. Ed. 50*, 5024

# Formylglycine generating enzyme - based tagging



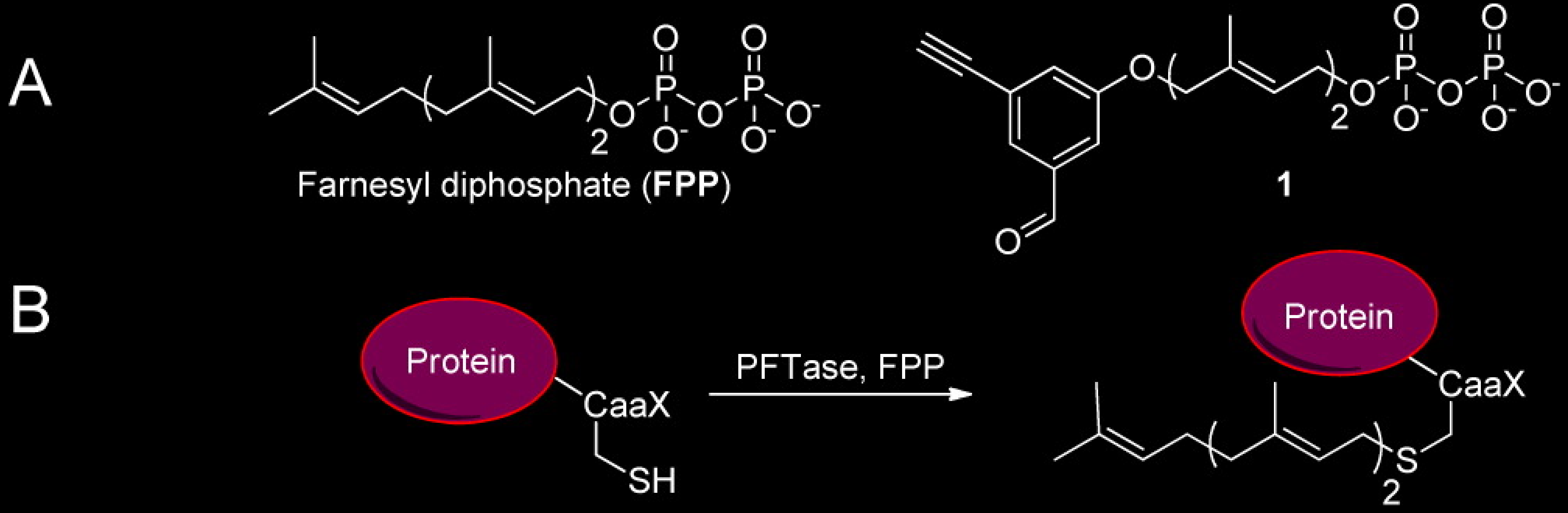
- Hexapeptide based consensus sequence
- Formylglycine generating enzyme



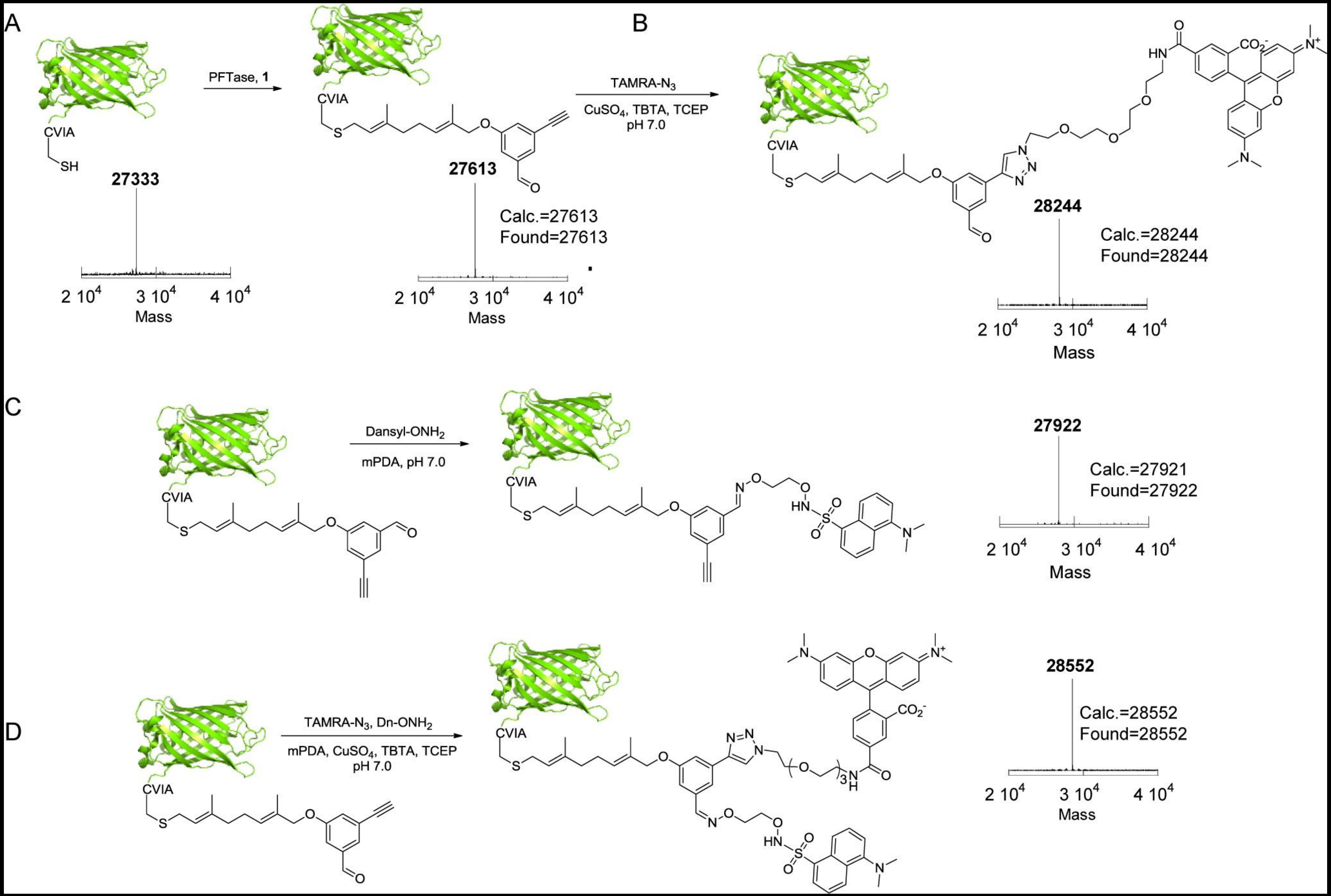

P. Wu et al. (2009). *Proc. Natl. Acad. Sci. U.S.A. 106*, 3000

# Protein farnesyl transferase

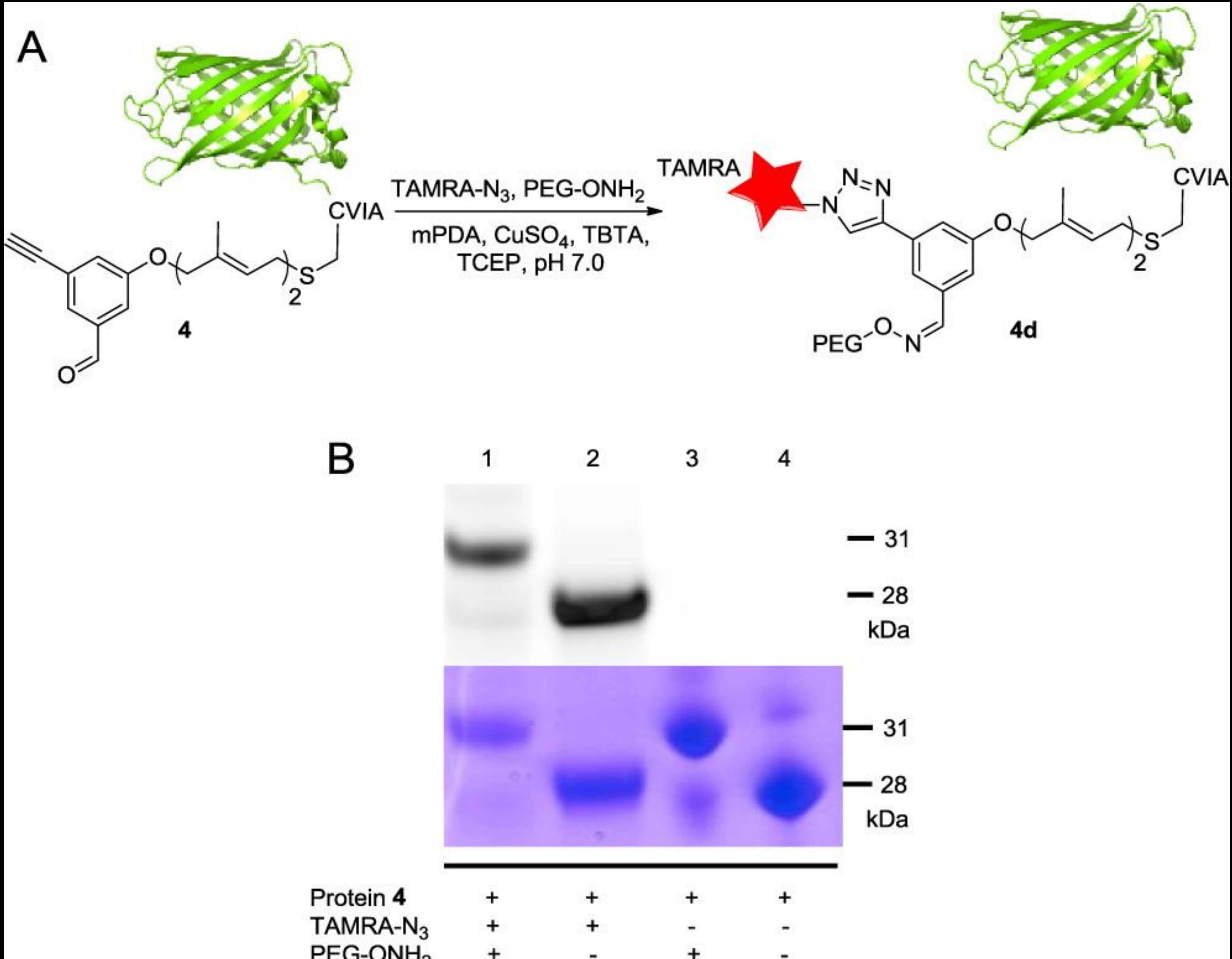
- CaaX box at the C-terminus of a protein (CVIA)




P. Wu et al. (2009). *Proc. Natl. Acad. Sci. U.S.A. 106*, 3000




P. Wu et al. (2009). *Proc. Natl. Acad. Sci. U.S.A. 106*, 3000




P. Wu et al. (2009). *Proc. Natl. Acad. Sci. U.S.A. 106*, 3000

# Protein Tags – fusion proteins

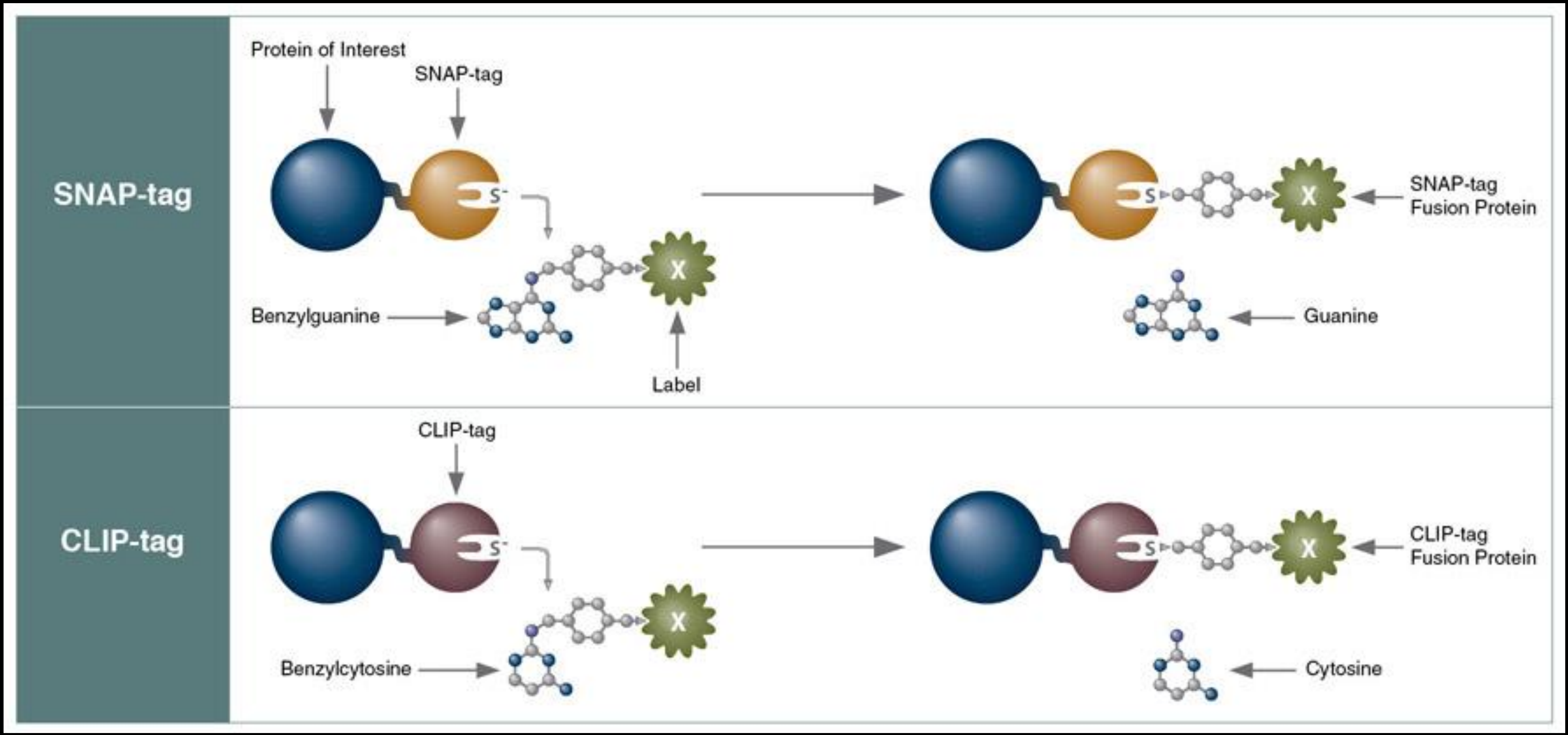
- Peptide (small protein) sequences with unique activity
- Can be added by recombinant DNA techniques
- Fluorescent tags (GFPs)
- Anchoring tags (His6)
- Labeling tags (Cys4)
- Enzymatic function for labeling (SNAP, Clip, Halo tags)

D. M. Patterson et al. (2014) *ACS Chem. Biol. 9*, 592

# Snap and Clip tags



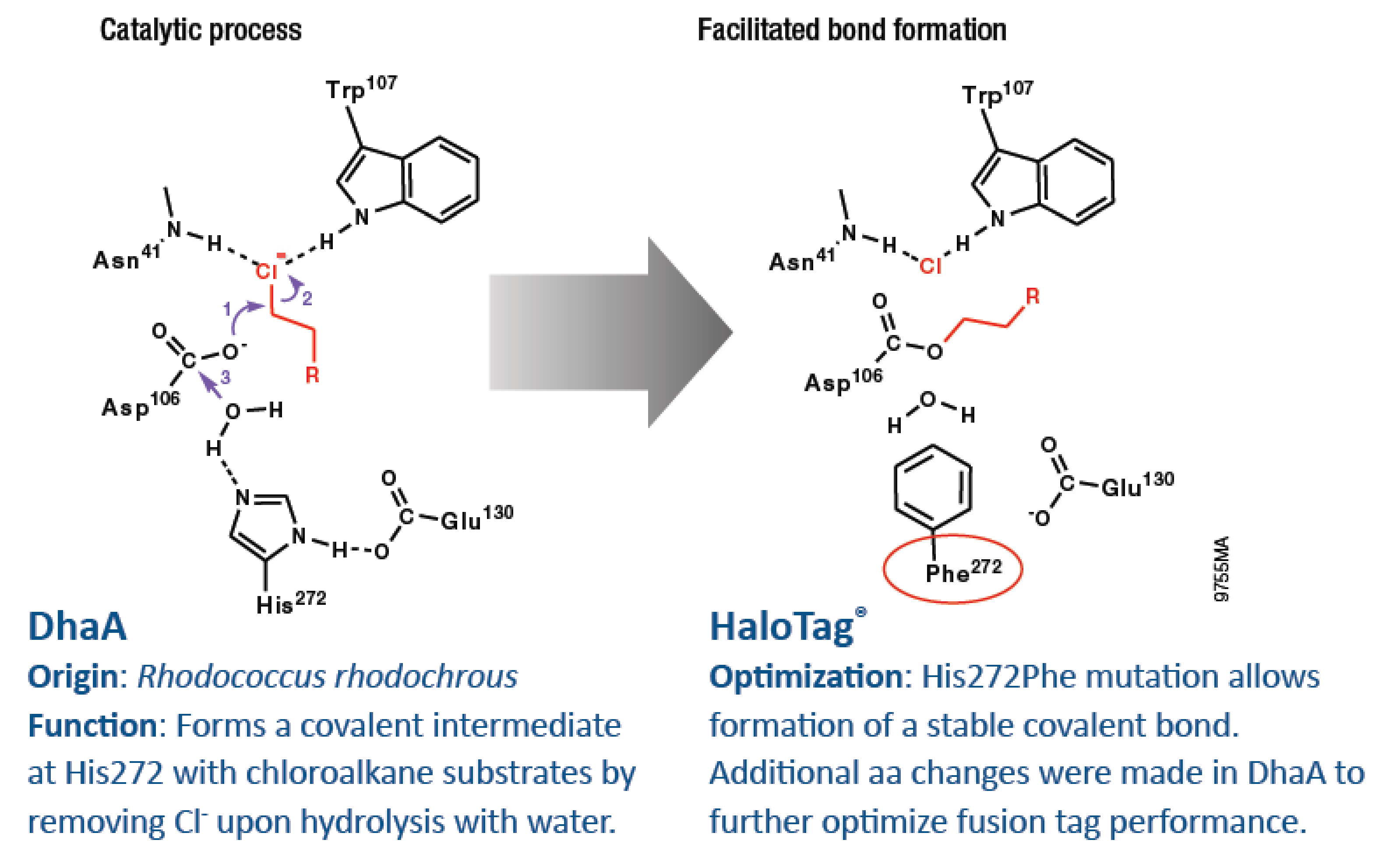
## O-6-Methylguanine-DNA methyltransferase activity




- A. Juillerat A, et al.(2003). *Chem. Biol. 10,* 313

# Halo tag

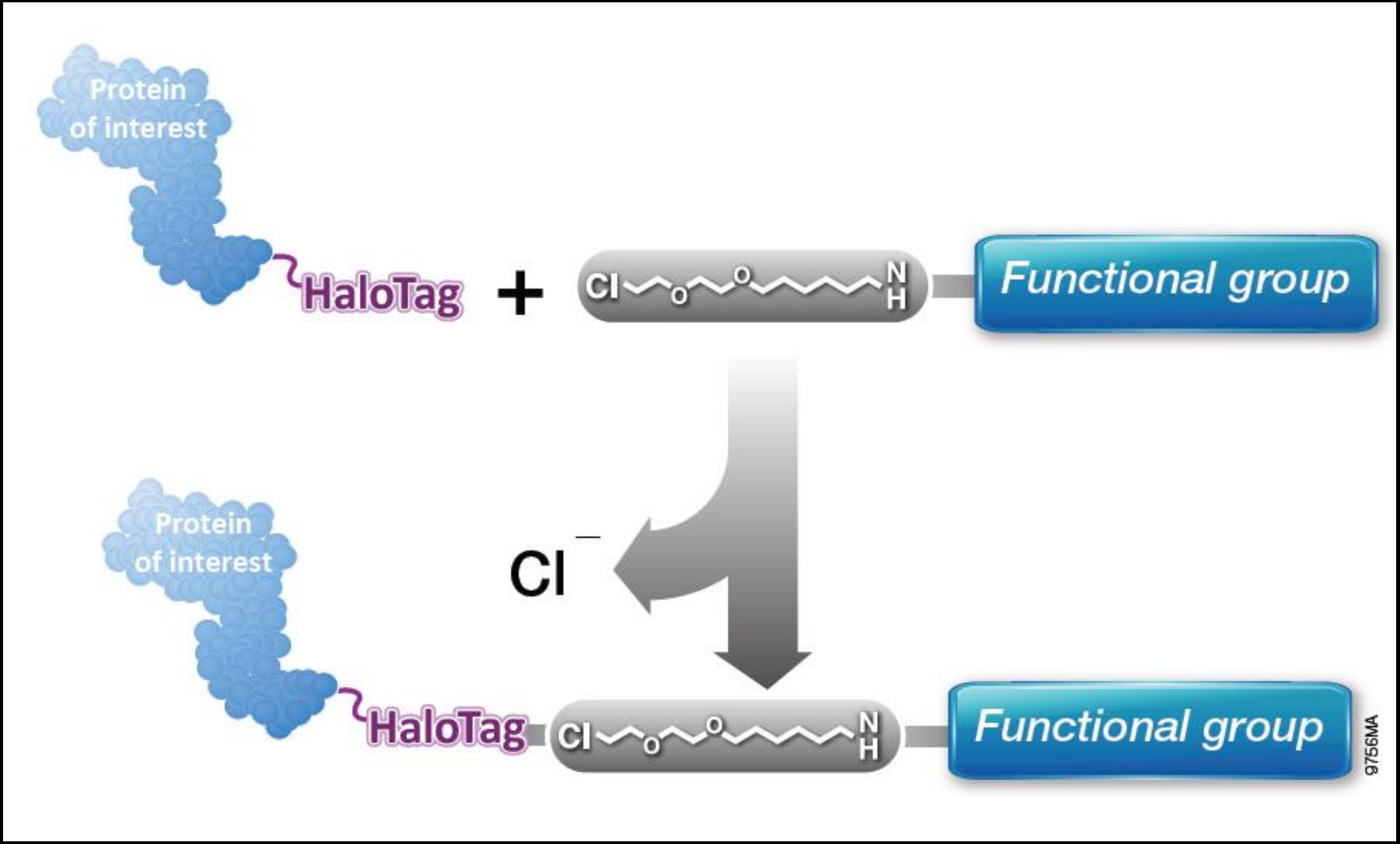





- G. V. Los et al. (2008) *ACS Chem. Biol.*, *3*, 373
- www.promega.co.uk

# Halo tag

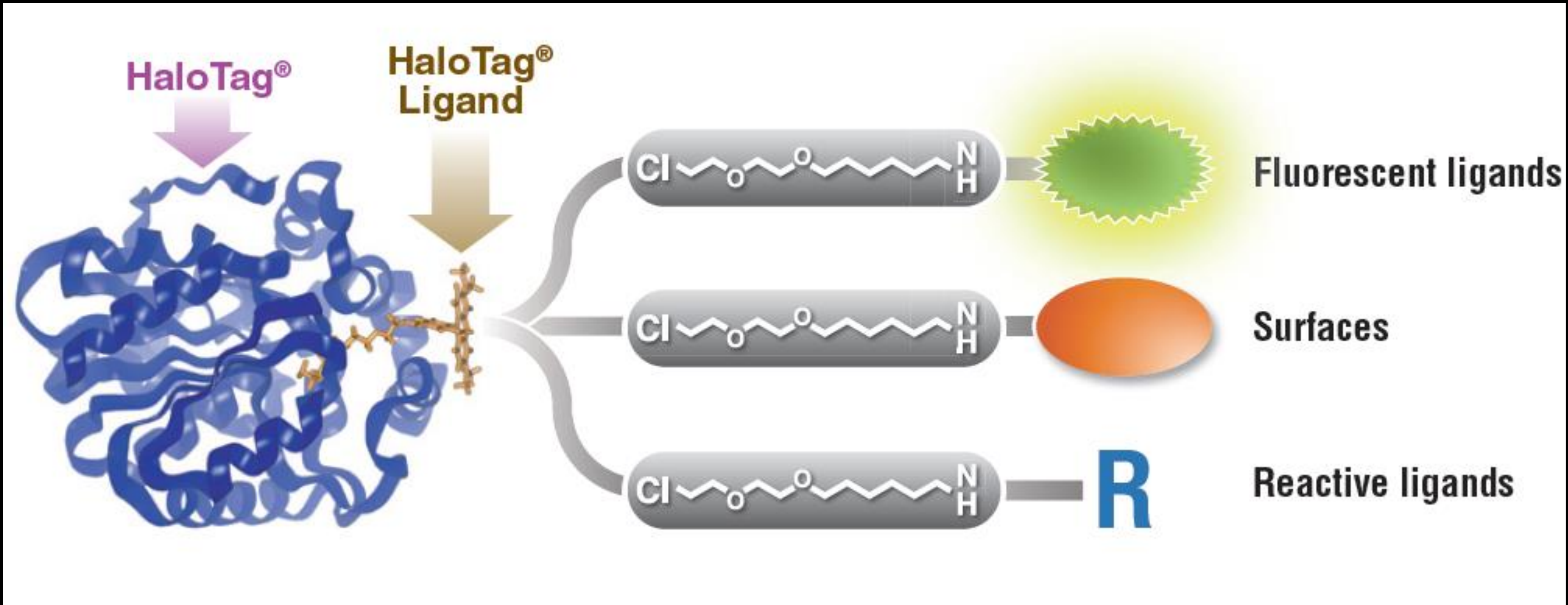





- www.promega.co.uk

# Halo tag

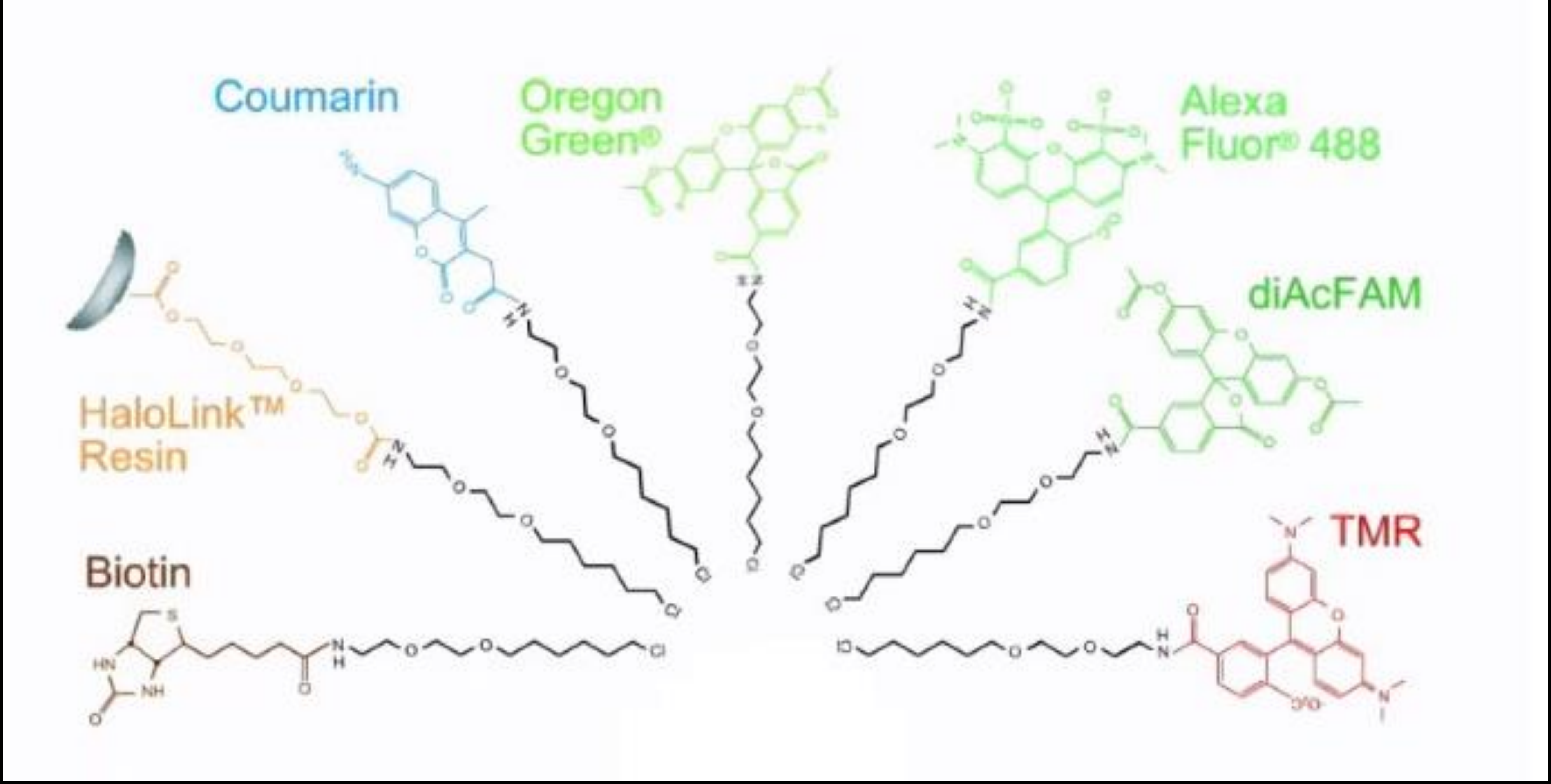





- www.promega.co.uk

# Halo tag



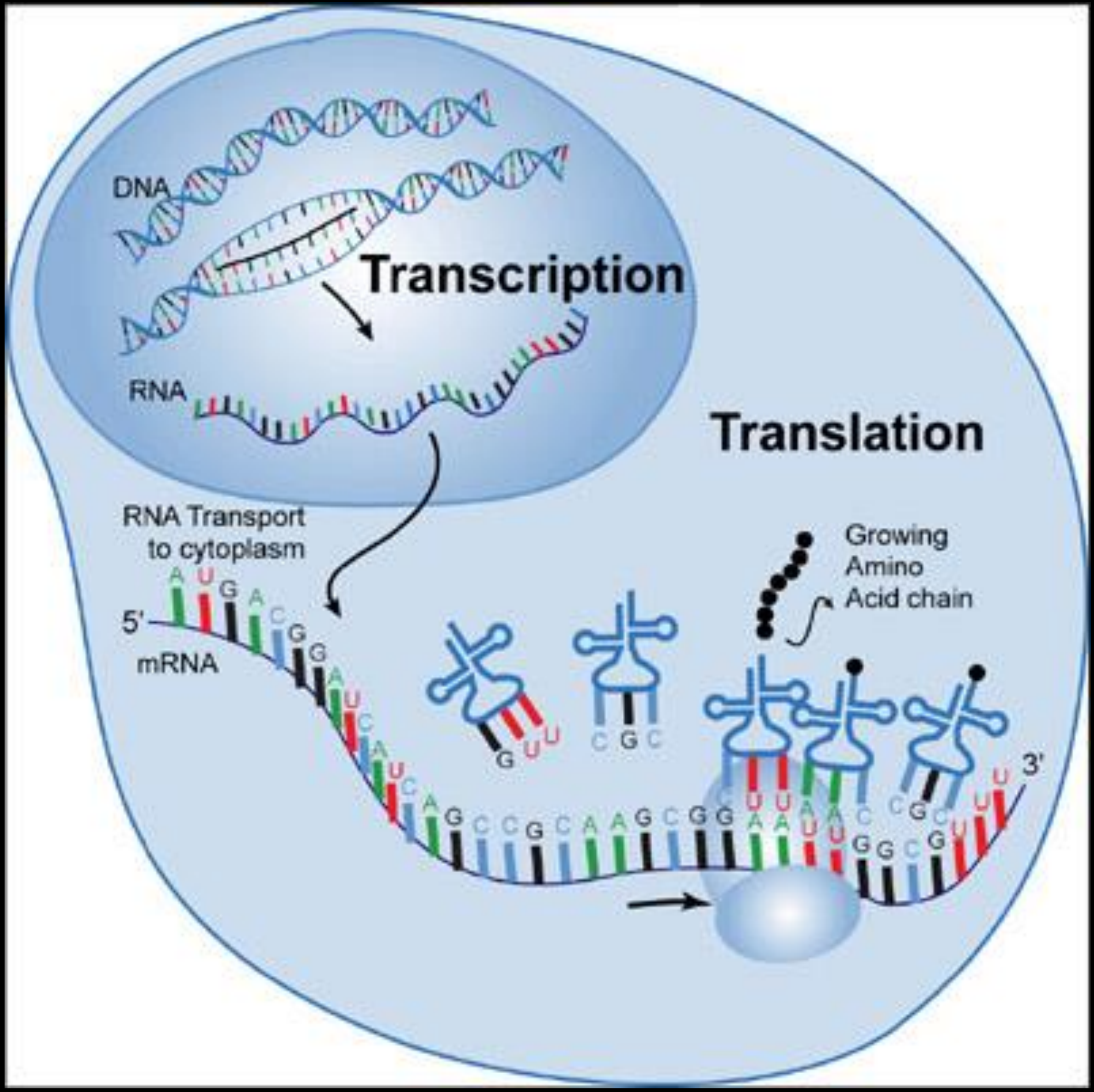



- G. V. Los et al. (2008) *ACS Chem. Biol.*, *3*, 373
- www.promega.co.uk

# 3.3. Genetic encoding of unnatural amino acids



- One AA = 3 bases
- DNA triplets
- mRNA codons
- tRNA anticodons



# Genetic code



- 64 possible triplets
- 20 (22) natural AAs encoded directly
- More triplets than Aas
- Genetic code is degenerate
- First two bases are fixed
- Third is loose (wobbling)
- Met, Trp egy kodon
- One START codon
- Three STOP codons

Second letter

| First letter (5′ end) | U | C | A | G | Third letter (3′ end) |
|---|---|---|---|---|---|
| U | UUU Phe | UCU Ser | UAU Tyr | UGU Cys | U |
| | UUC Phe | UCC Ser | UAC Tyr | UGC Cys | C |
| | UUA Leu | UCA Ser | UAA Stop | UGA Stop | A |
| | UUG Leu | UCG Ser | UAG Stop | UGG Trp | G |
| C | CUU Leu | CCU Pro | CAU His | CGU Arg | U |
| | CUC Leu | CCC Pro | CAC His | CGC Arg | C |
| | CUA Leu | CCA Pro | CAA Gln | CGA Arg | A |
| | CUG Leu | CCG Pro | CAG Gln | CGG Arg | G |
| A | AUU Ile | ACU Thr | AAU Asn | AGU Ser | U |
| | AUC Ile | ACC Thr | AAC Asn | AGC Ser | C |
| | AUA Ile | ACA Thr | AAA Lys | AGA Arg | A |
| | AUG Met or start | ACG Thr | AAG Lys | AGG Arg | G |
| G | GUU Val | GCU Ala | GAU Asp | GGU Gly | U |
| | GUC Val | GCC Ala | GAC Asp | GGC Gly | C |
| | GUA Val | GCA Ala | GAA Glu | GGA Gly | A |
| | GUG Val | GCG Ala | GAG Glu | GGG Gly | G |

= Stop codon

= Start codon

# STOP codons

- Amber (UAG), Ochre (UAA) and Opal (UGA)
- Amber is very rare
- No tRNA exists to bind them
- In certain bacteria (Archia) Pyl is encoded on UGA
- With Pyl tRNA and Pyl-tRNA synthase Pyl can be encoded in other organisms


- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Expanding the genetic code



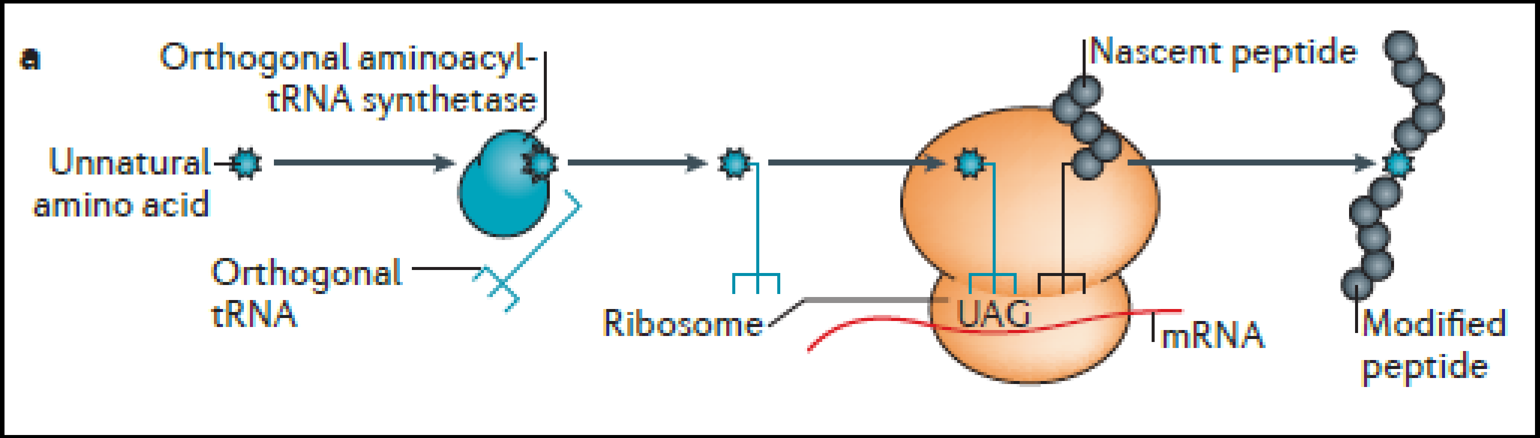
- With Pyl tRNA and Pyl-tRNA synthase Pyl can be encoded in other organisms
- tRNA / RS pair that acylates one AA to one tRNA without „cross-talk” is orthogonal
- tRNA is not a substrate of normal synthetases
- Synthetase that does not aminoacylate normal tRNAs


- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Expanding the genetic code



- AMBER supression by orthogonal AARS / tRNA
- AARS / tRNA pair that acylates one AA to one tRNA without „cross-talk” is orthogonal
- AARS / tRNA pairs that accept UAAs

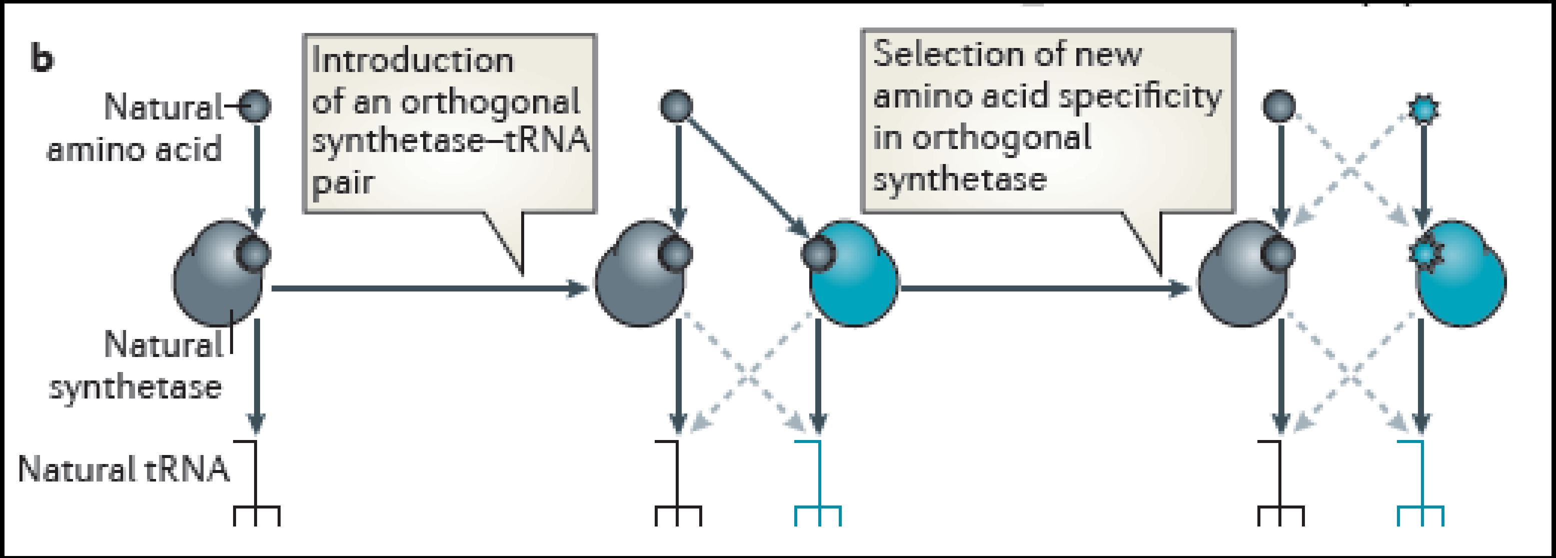




- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Generation of orthogonal AARS / tRNA

- Selection of a AARS / tRNA from an organism that evolutionariliy different from the host (ea. Recognizes UGA)

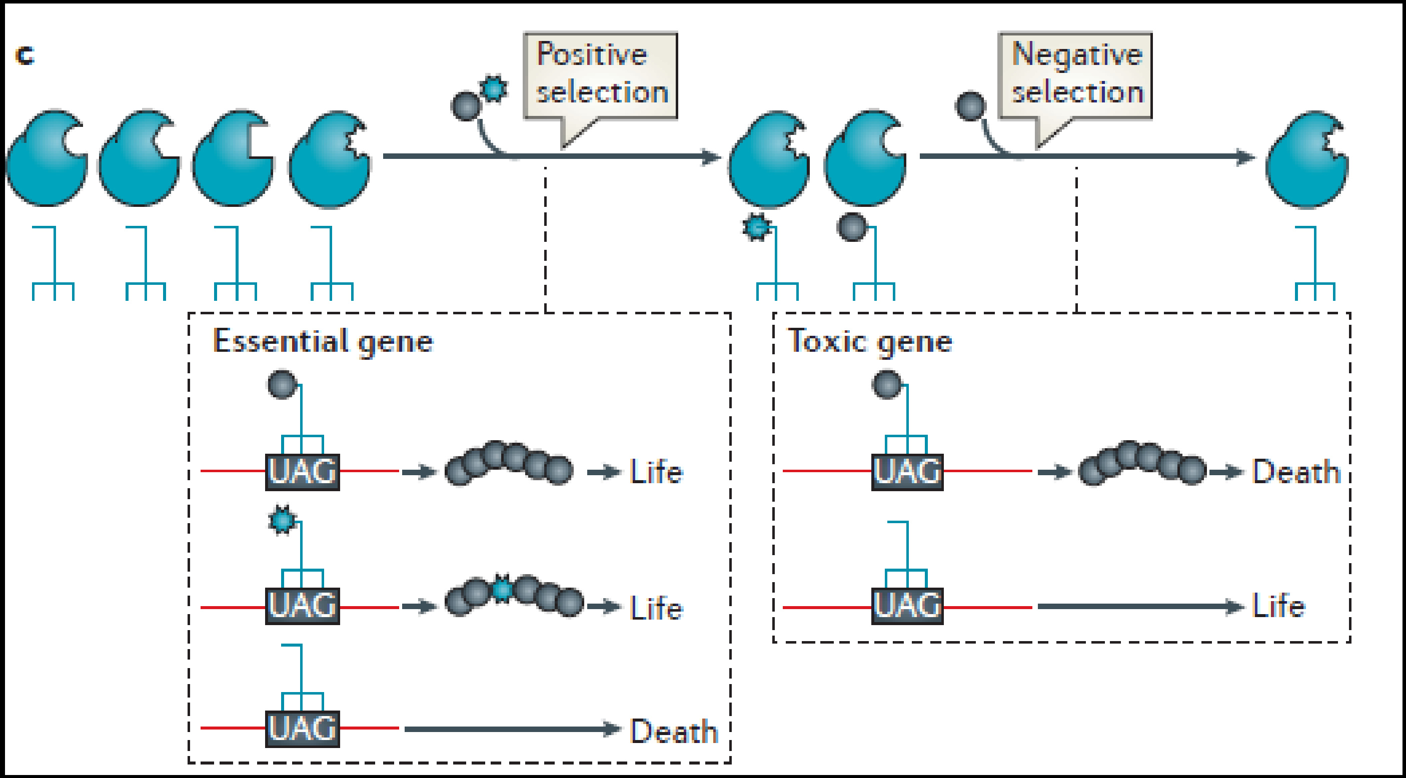



- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Generation of orthogonal AARS / tRNA



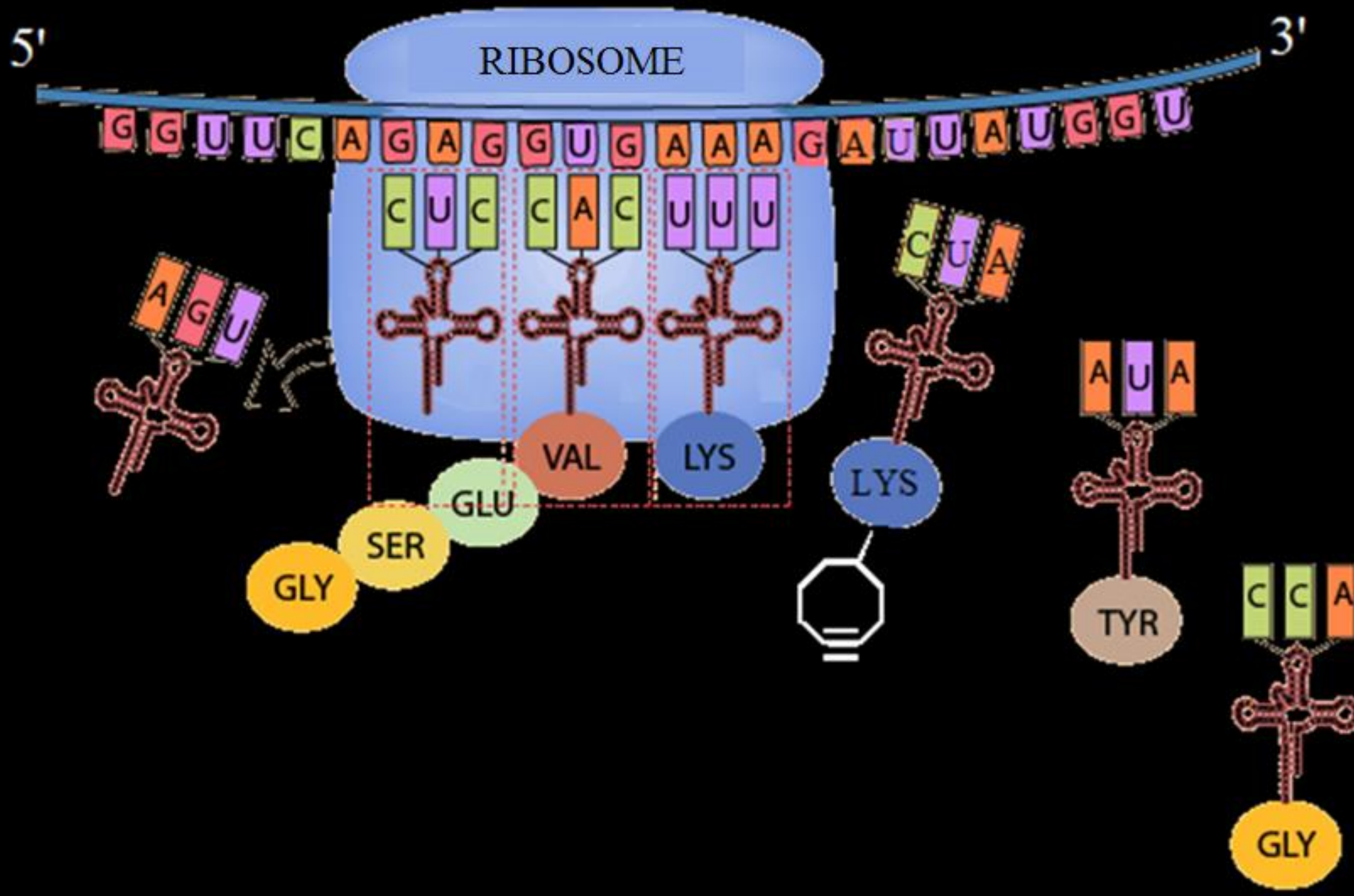
- As AARS in general still recognizes natural Aas the enzyme is mutated (library) then RS that is orthogonal is selected





- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Genetic incorporation of chemical reporter




T. Plass, S. Milles, C. Koehler, C. Schultz, E. A. Lemke (2011) *Angew. Chem. Int. Ed*. *50,* 3878

# Generation of orthogonal AARS / tRNA



- Essentially four AARS / tRNA is used as starting point
- Methanococcus jannaschii TyrRS / $tRNA_{CUA}$
- E. coli TyrRS / $tRNA_{CUA}$ EcTyrRS / $tRNA_{CUA}$
- EcLeuRS / $tRNA_{CUA}$
- Methanosarcina sp. PylRS / $tRNA_{CUA}$
  - Does not use AA of the 20 canonical
  - Can be selected in E. coli


- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Implementation of AARS / tRNA

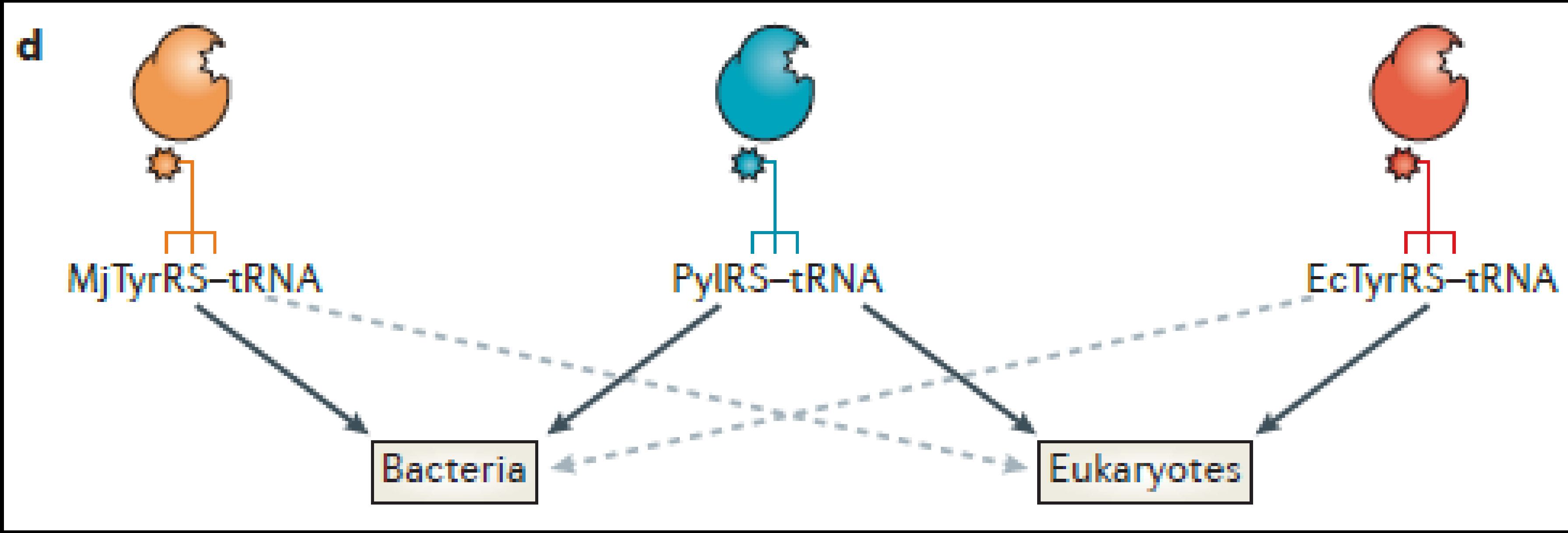


- Old fashioned way: pre-loaded tRNA-UAA
- Later AARS and tRNA coding genes were implemented into host genome (plasmid)
- UAG modified protein of interest is also implemented into plasmids (PCR, recombinant DNA techniques)
- UAA is added as nutrient


- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

# Generation of orthogonal AARS / tRNA

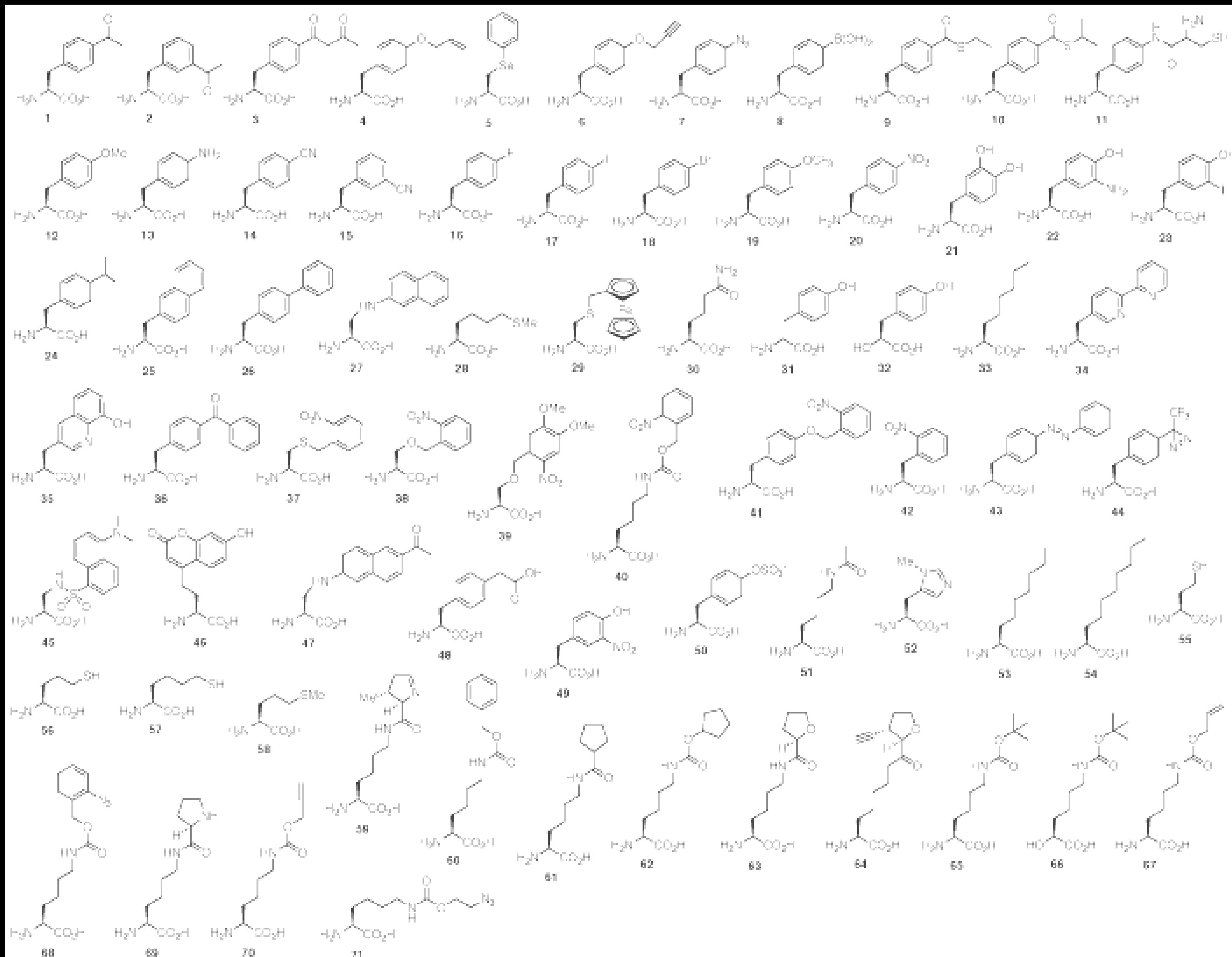





- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

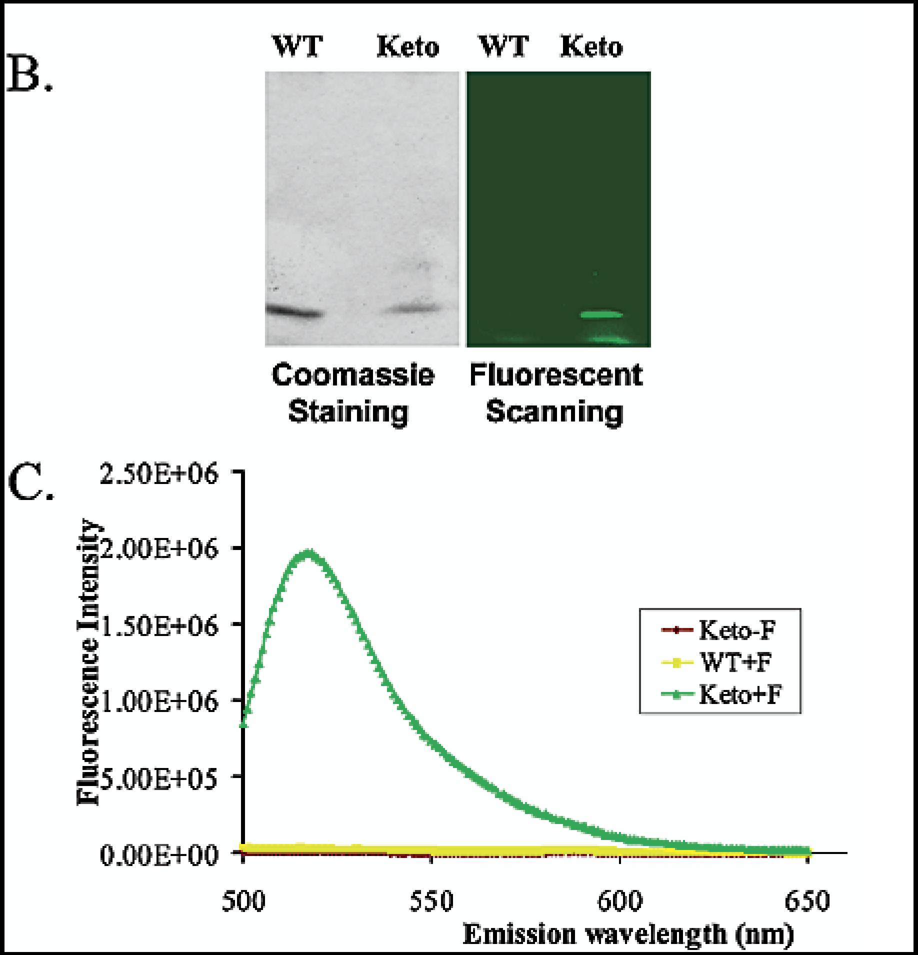
# Encoded UAAs

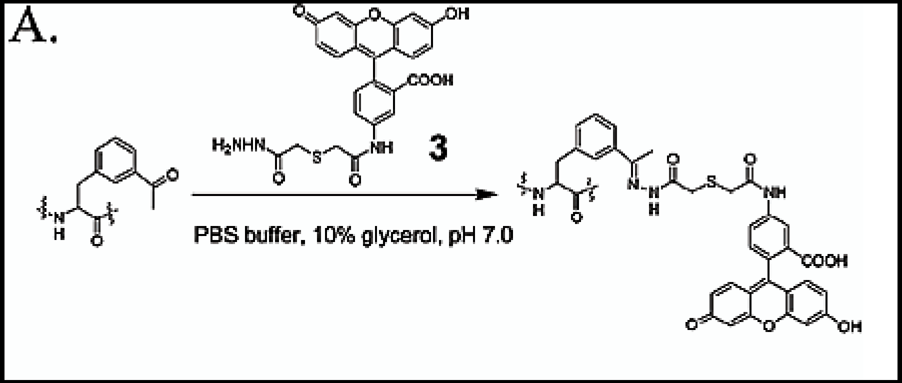





- C. C. Liu, P. G. Schultz ( 2010) *Annu. Rev. Biochem. 79*, 413
- L. Davis, J. W. Chin ( 2012) *Nat. Rev. Mol. Cell Biol*. 13, 168

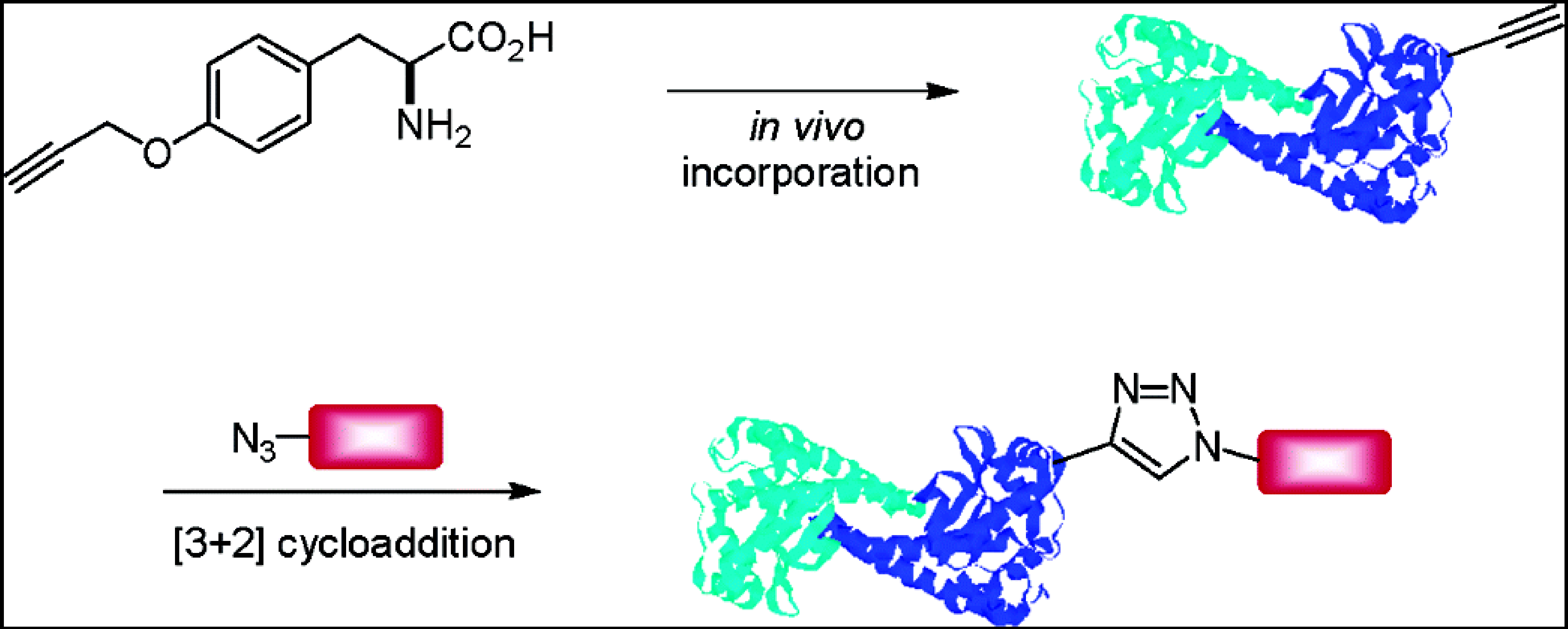
# Encoded UAAs



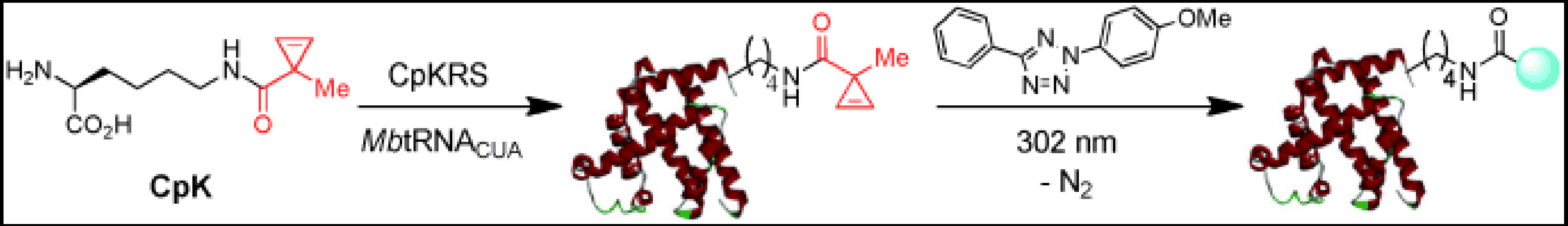

• Zhang, Z., Smith, B. A., Wang, L., Brock, A., Cho, C., and Schultz, P. G. ( 2003) *Biochemistry 42*, 6735

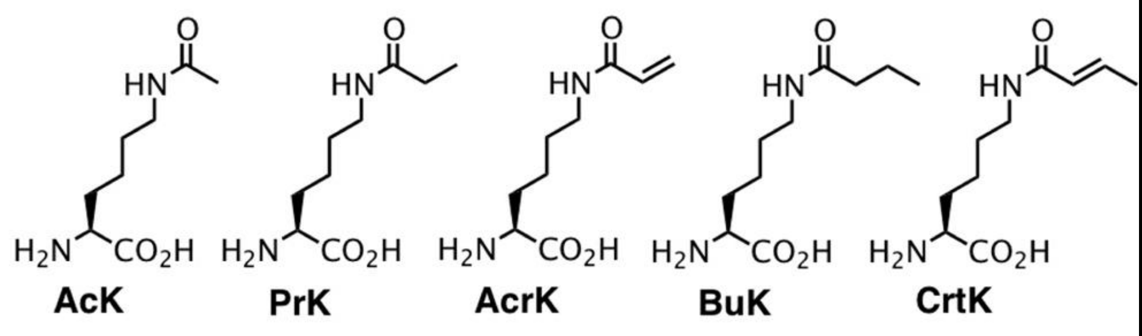
# Encoded UAAs

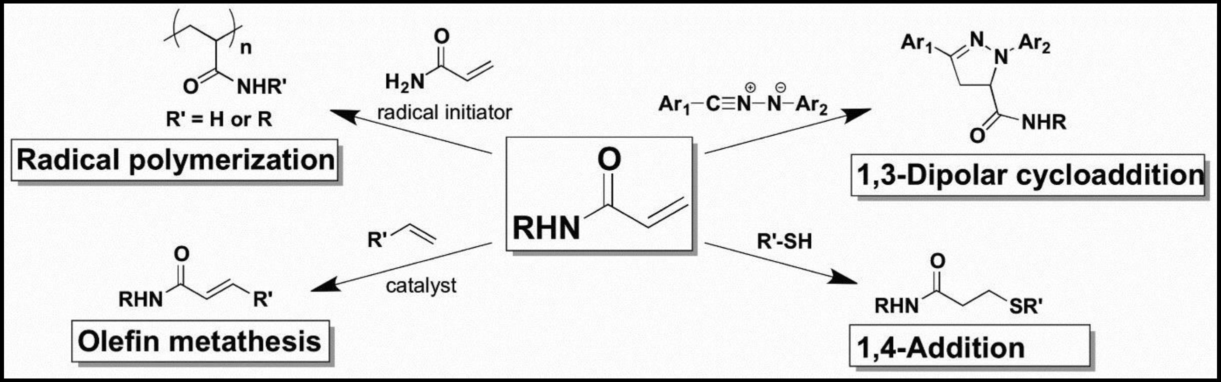



- Deiters, A., Cropp, T. A., Mukherji, M., Chin, J. W., Anderson, J. C., and Schultz, P. G. ( 2003) *J. Am. Chem. Soc. 125*, 11782

# Encoded UAAs



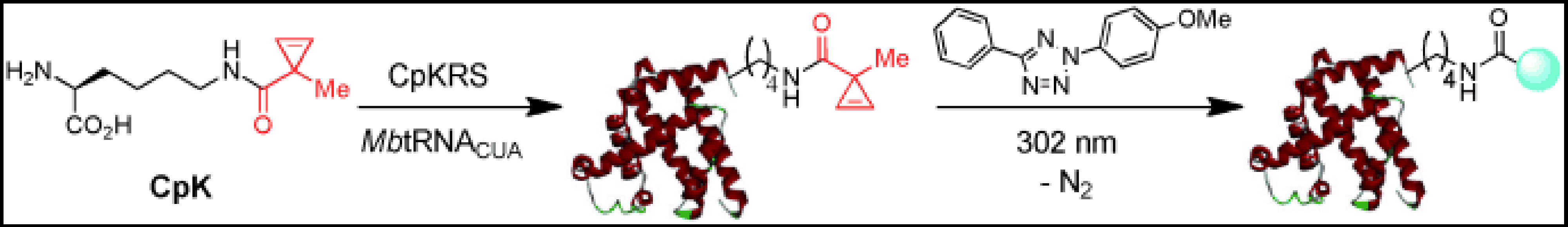



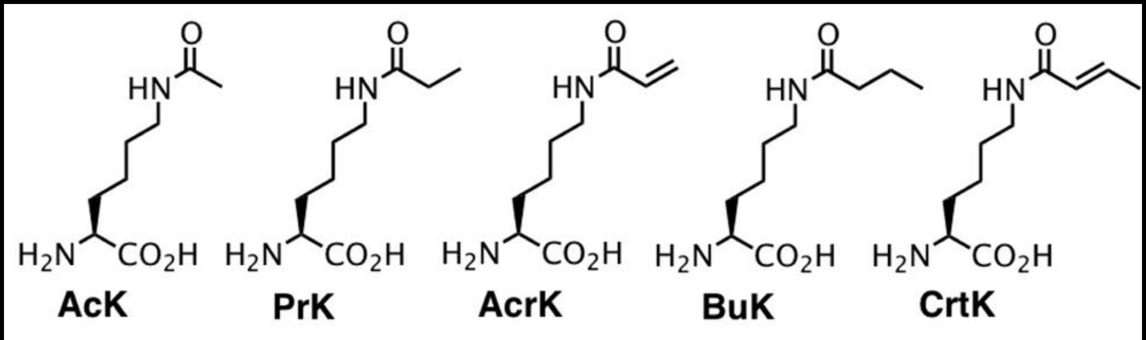



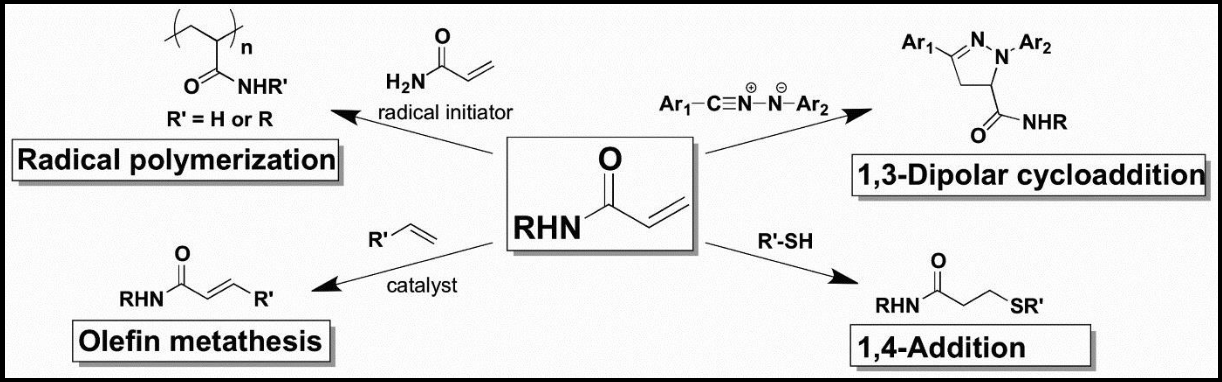

- Yu, Z., Pan, Y., Wang, Z., Wang, J., and Lin, Q. ( 2012) Angew. Chem., Int. Ed. 51, 10600
- Lee, Y. J., Wu, B., Raymond, J. E., Zeng, Y., Fang, X., Wooley, K. L., and Liu, W. R. ( 2013) ACS Chem. Biol. 8, 1664

# Encoded UAAs





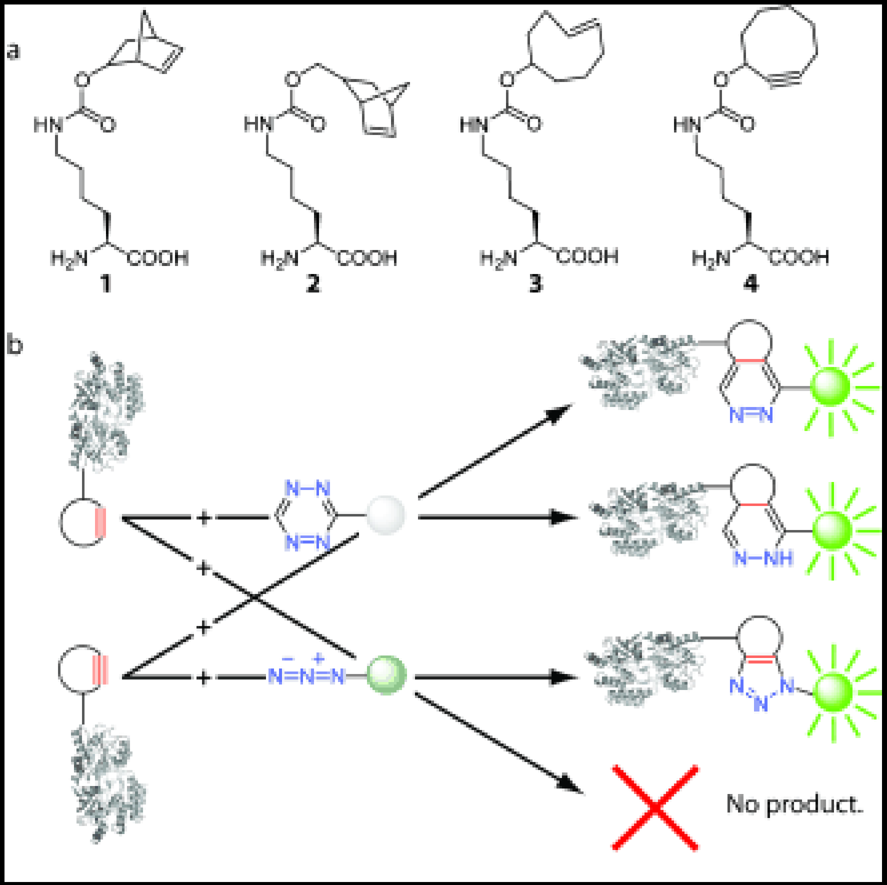



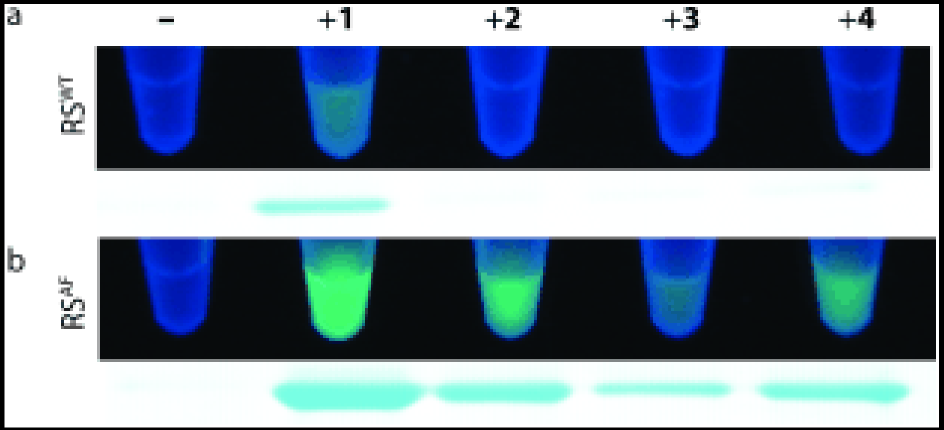


- Yu, Z., Pan, Y., Wang, Z., Wang, J., and Lin, Q. ( 2012) Angew. Chem., Int. Ed. 51, 10600
- Lee, Y. J., Wu, B., Raymond, J. E., Zeng, Y., Fang, X., Wooley, K. L., and Liu, W. R. ( 2013) ACS Chem. Biol. 8, 1664

# Encoded UAAs



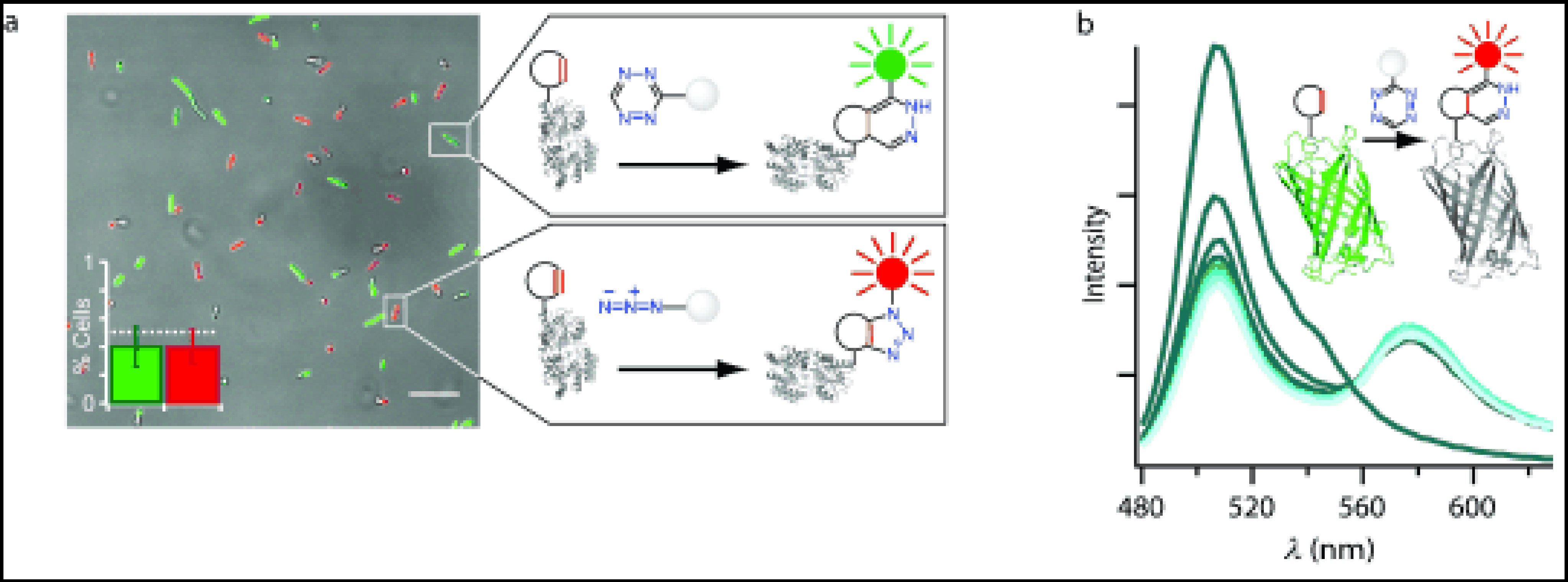



- Plass, T., Milles, S., Koehler, C., Szymanski, J., Mueller, R., Wiessler, M., Schultz, C., and Lemke, E. A. ( 2012) Angew. Chem., Int. Ed. 51, 4166

# Encoded UAAs



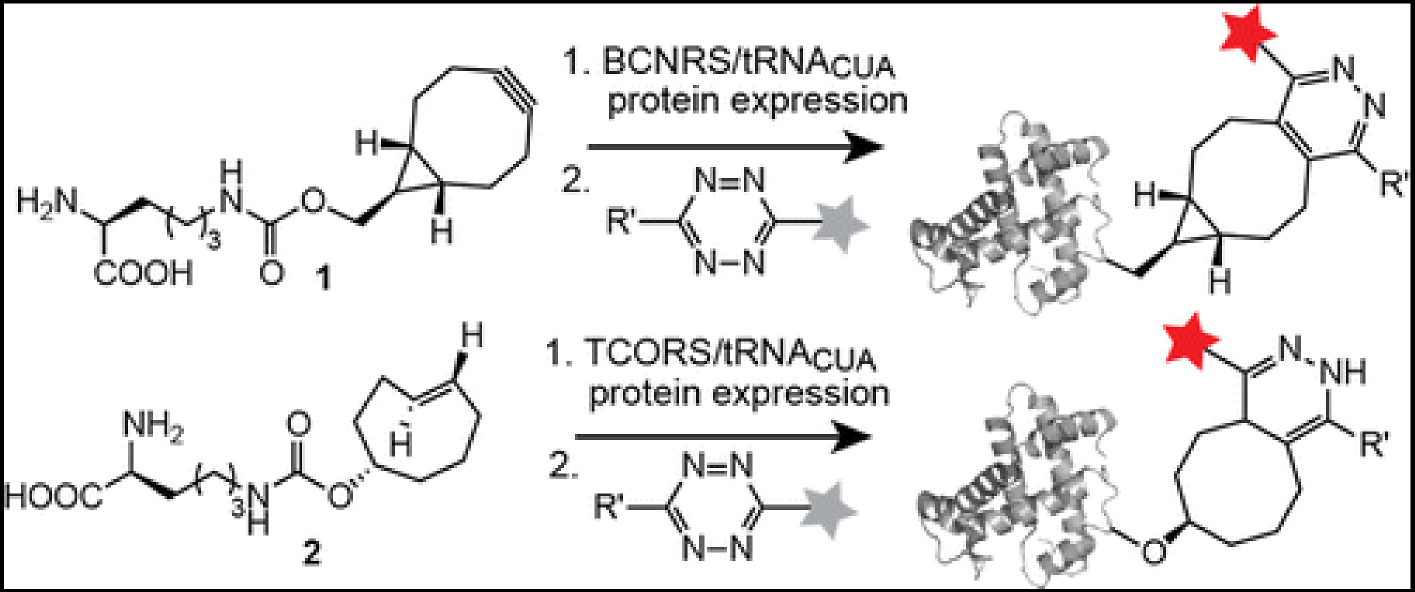




- Plass, T., Milles, S., Koehler, C., Szymanski, J., Mueller, R., Wiessler, M., Schultz, C., and Lemke, E. A. ( 2012) Angew. Chem., Int. Ed. 51, 4166

# Encoded UAAs






- Lang, K., Davis, L., Wallace, S., Mahesh, M., Cox, D. J., Blackman, M. L., Fox, J. M., and Chin, J. W. ( 2012) J. Am. Chem. Soc. 134, 10317

# Further examples

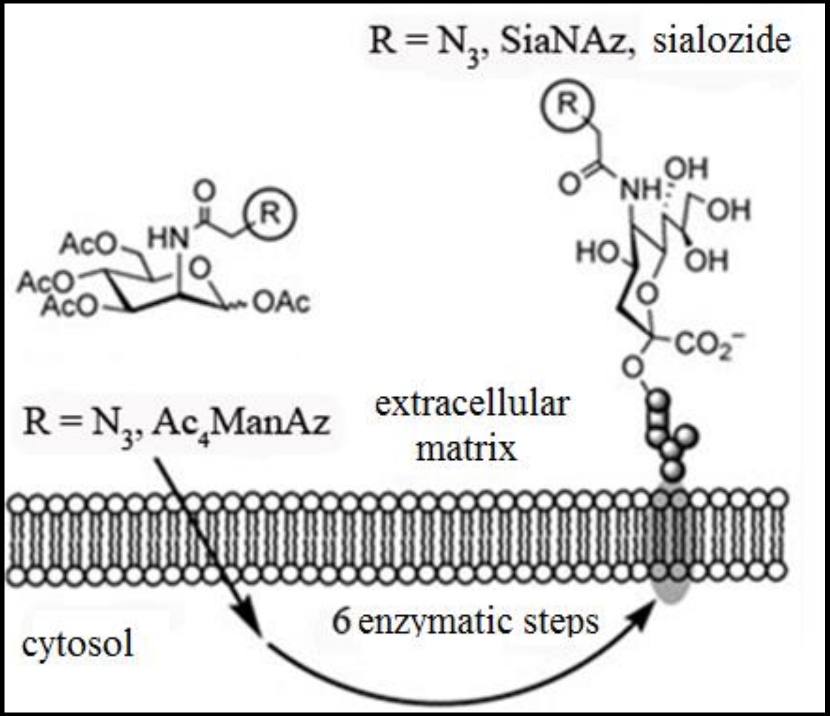


- Tetrazine, tetrazole encoding
- Quadrupole encoding (more challenging)
  - Ribosome has to be altered (frame reading)
  - Multiple encoding in combination with amber supression

# Metabolic labeling

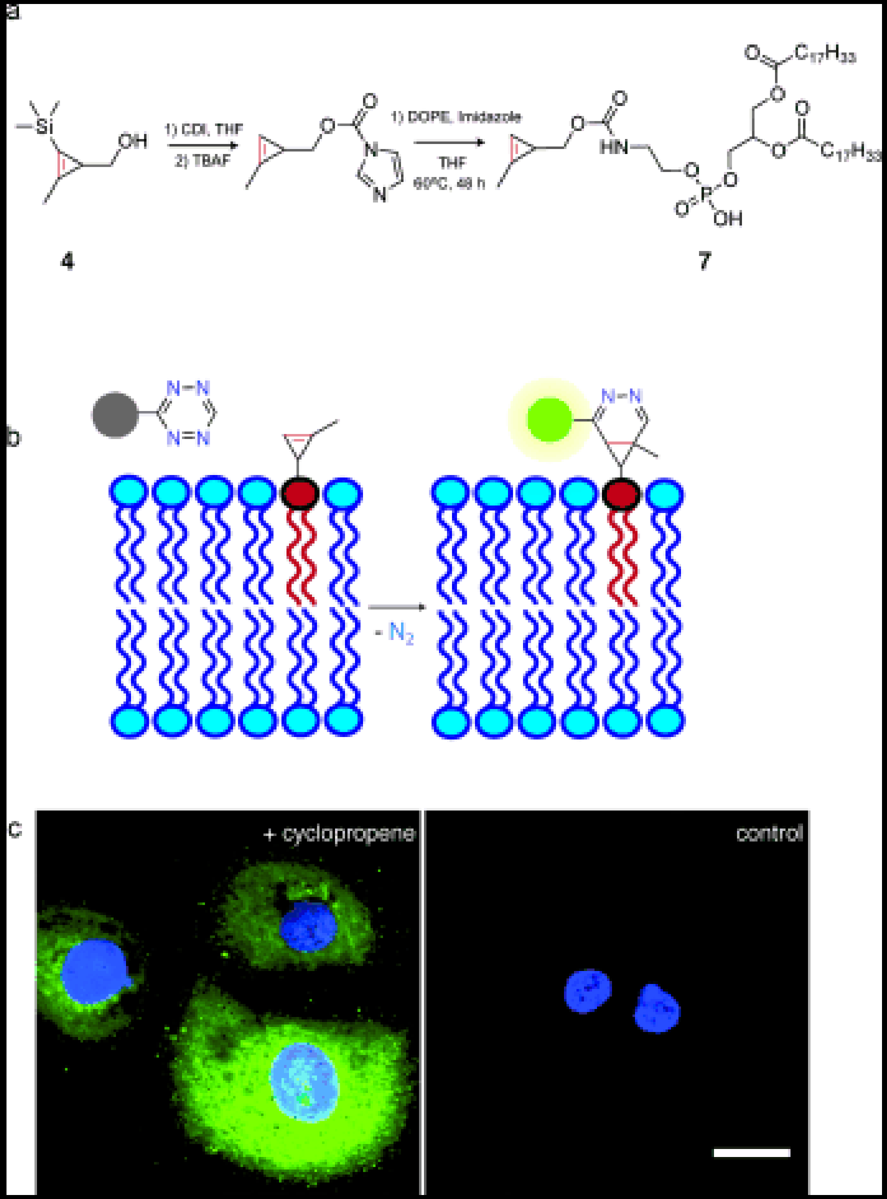


- ▶Modified metabolites (mainly sugars, azido, alkyne)
- ▶Only small modifications are allowed
- ▶Used for glycan modification (surface glycan)



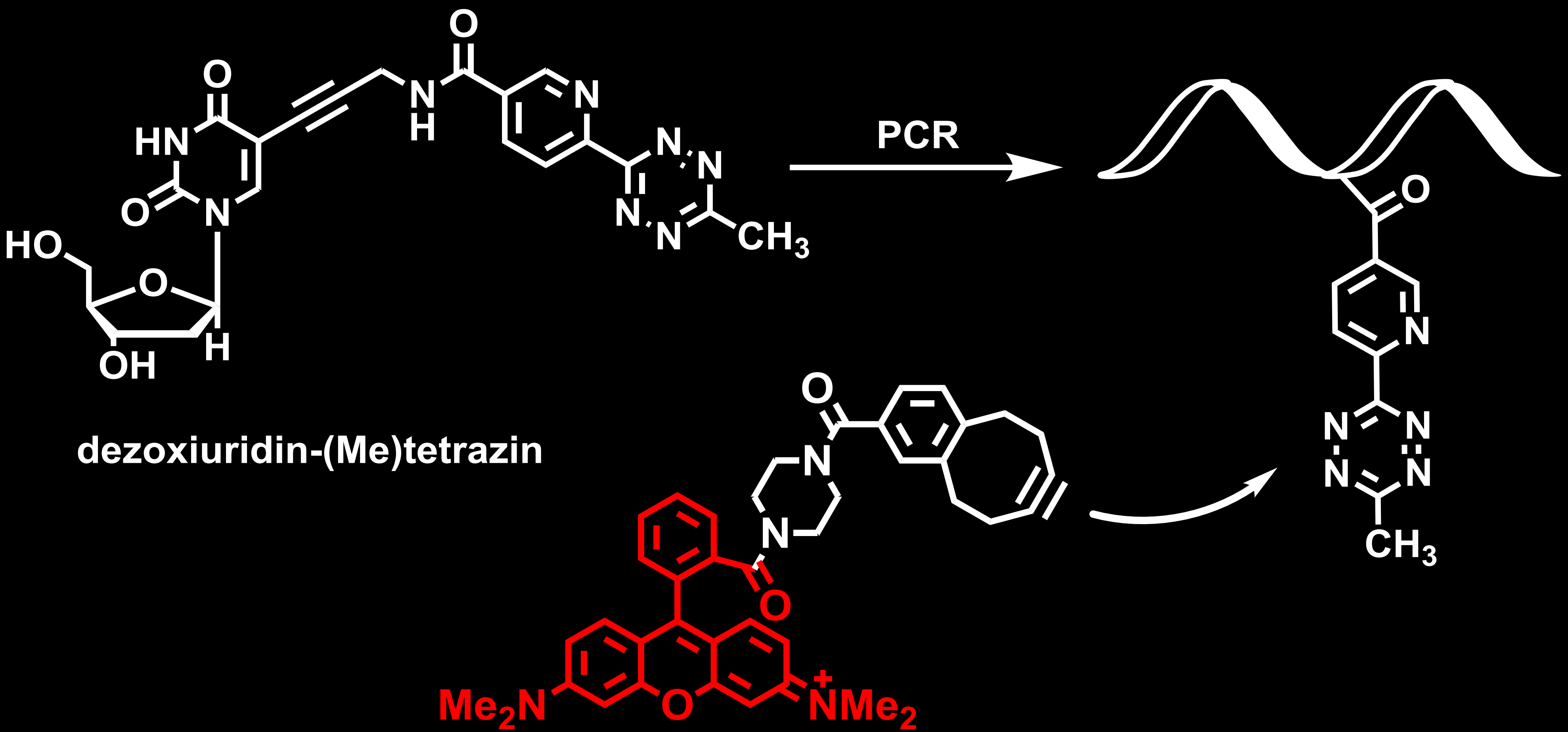

• Prescher, J. A. and Bertozzi, C. R. ( 2005) Nat. Chem. Biol. 1, 13

# Lipid labeling




- Yang, J., et al. (2012) Angew. Chem. Int. Ed., 51, 7476

# Nucleic acid labeling

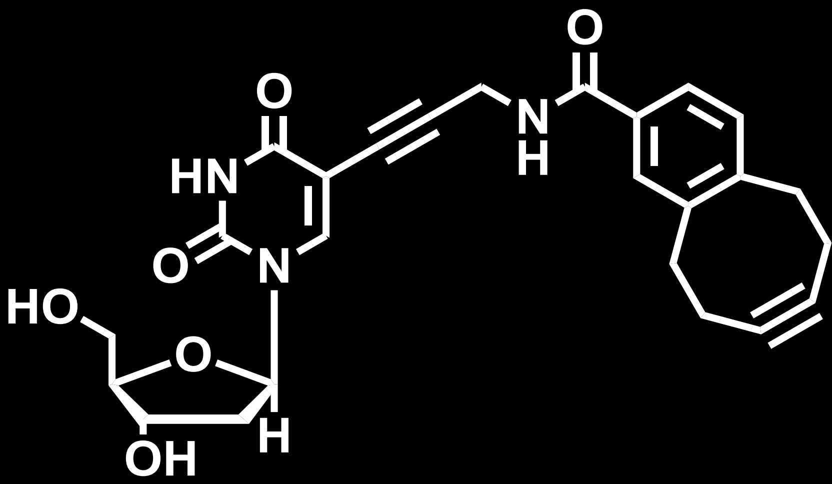



**dezoxiuridin-(Me)tetrazin**

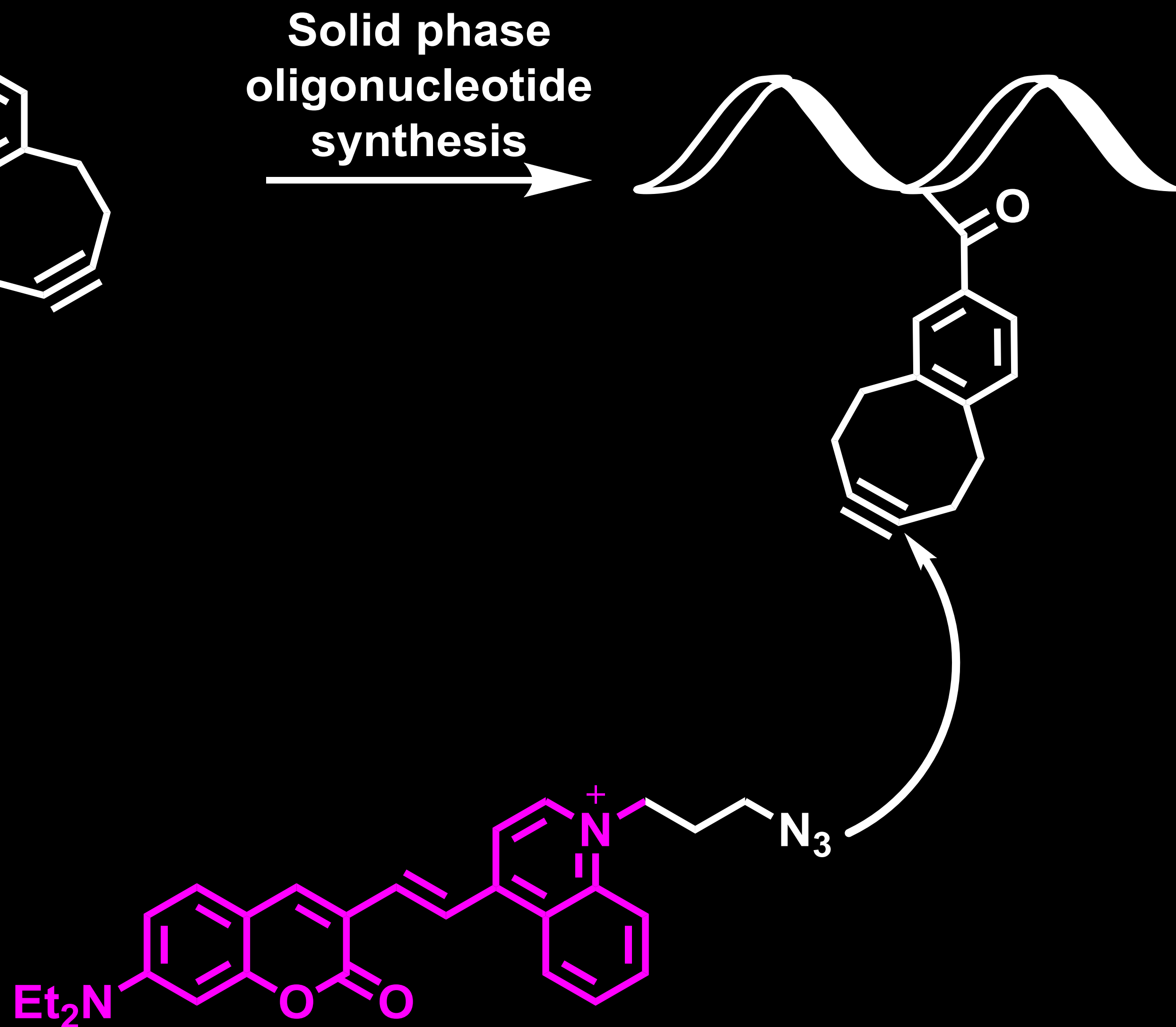
**COMBO-Rhod**

G. B. Cserép, O. Demeter, E. Bätzner, M. Kállay, H-A. Wagenknecht, P. Kele *In preparation*

# Nucleic acid labeling




**deoxyuridine-COMBO**




**$\lambda_{abs}$ = 575 nm**

**$\lambda_{em}$ = 725 nm**

---

- C. Stubnitzky, G. B. Cserép, B. R. Varga, P. Kele, H-A. Wagenknecht *In preparation*
- G. B. Cserép, K. Enyedi, A. Demeter, G. Mező, P. Kele (2013) *Chem. Asian. J. 8*, 494.

# Nucleic acid labeling