Chemical Manipulation of Protein Functions

-Utilizing incorporation of a genetically encoded unnatural amino acid-

Literature Seminar

20th Oct 2016

Kai Kitamura

Outline

1. Introduction
2. Incorporation of unnatural amino acids
3. Bioorthogonal ligand tethering (BOLT)
4. Protein-control using Staudinger reduction
   (Main Paper)
5. Summary

Introduction

■ Why is “protein-manipulation” required?

Biochemistry: traditionally focused on the study *in vitro*
(in a reconstituted system)

it cannot fully mimic a biological system

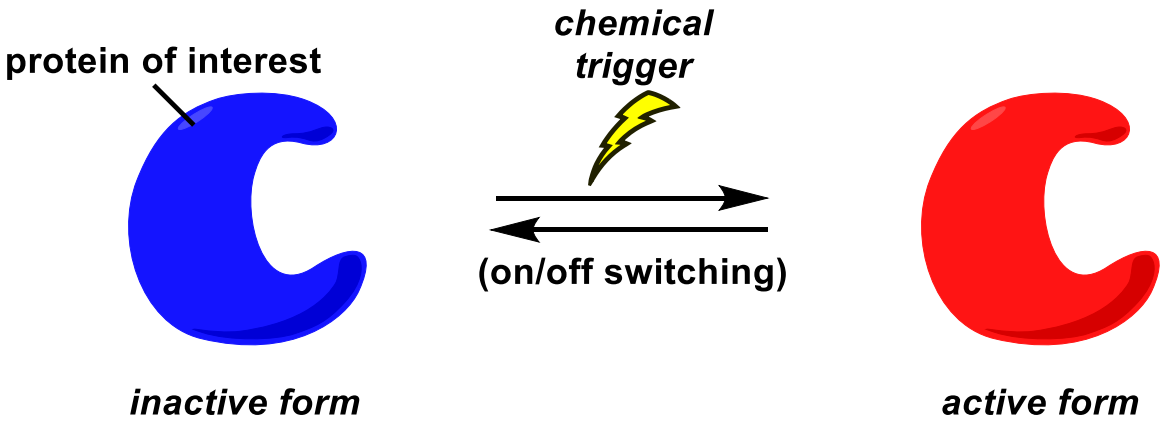
Factors such as:

- Dosage
- Intracellular Localization
- Timing
- Intermolecular Interactions

are important in living systems

Hahn, M.E.; Muir, T. W. *TRENDS in Biochem. Sci.* **2005**, 30, 1

Introduction



■ What is the *ideal* "protein-manipulation" method?

- immediate on/off switching of protein functions
- high selectivity of inhibition/activation
  (even between very similar isozymes)
- *in cell (in vivo)* applicability
- usability of the method
- versatile applicability to various proteins

<u>Rationally designed chemistry-driven strategies have appeared</u>

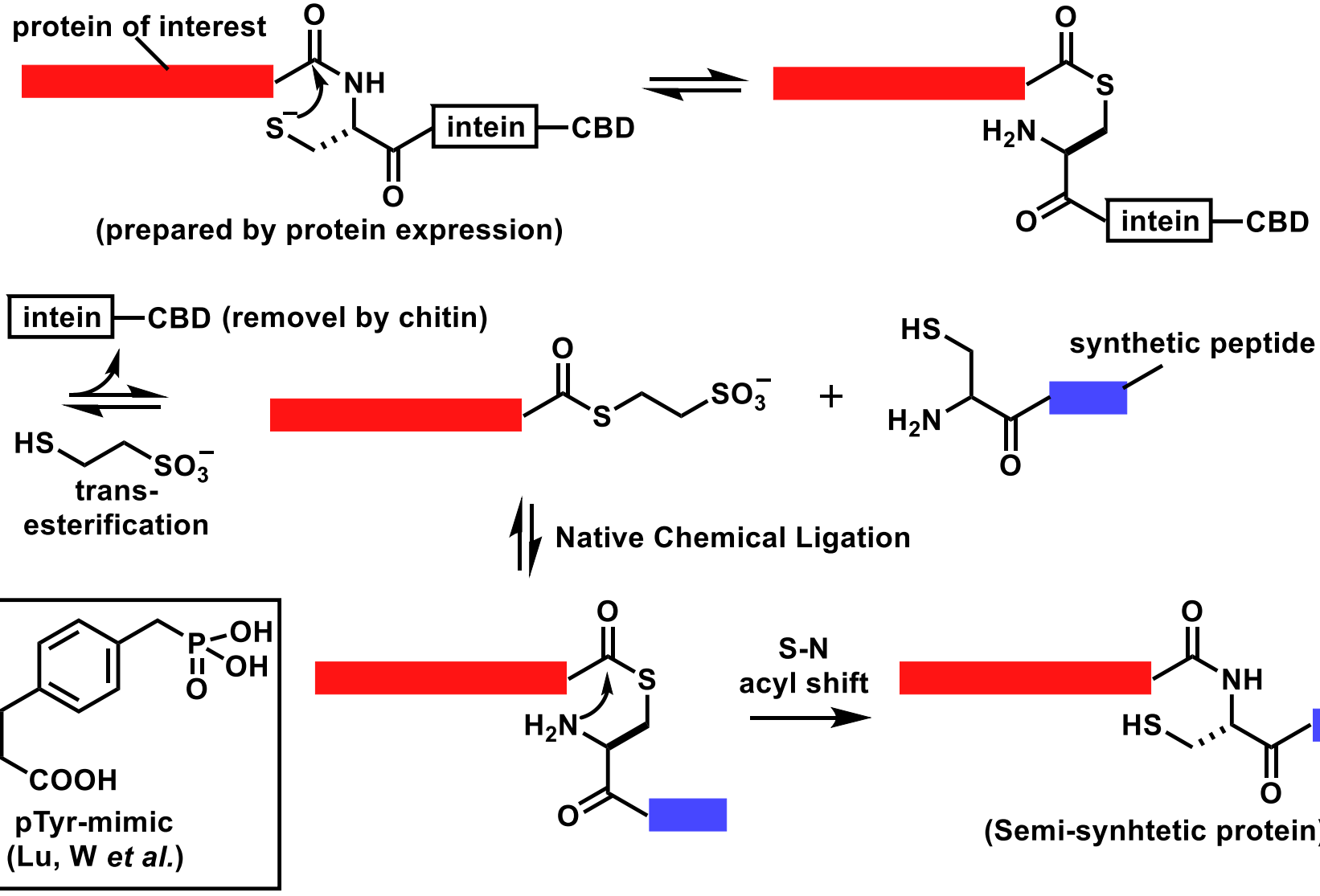
How to Control Proteins with Chemistry

1. (Inhibitors)
2. Expressed Protein Ligation (EPL)
3. Conditional Protein Splicing
4. Conditional Protein Degradation
5. Bump-and-Hole Strategy
6. Site-directed Unnatural Amino Acid Mutagenesis

Protein engineering techniques are indispensable

Intein-based Semi-synthesis (EPL: Expressed Protein Ligation)

Intein: Intron in protein processing/ **CBD**: Chitin-binding domain

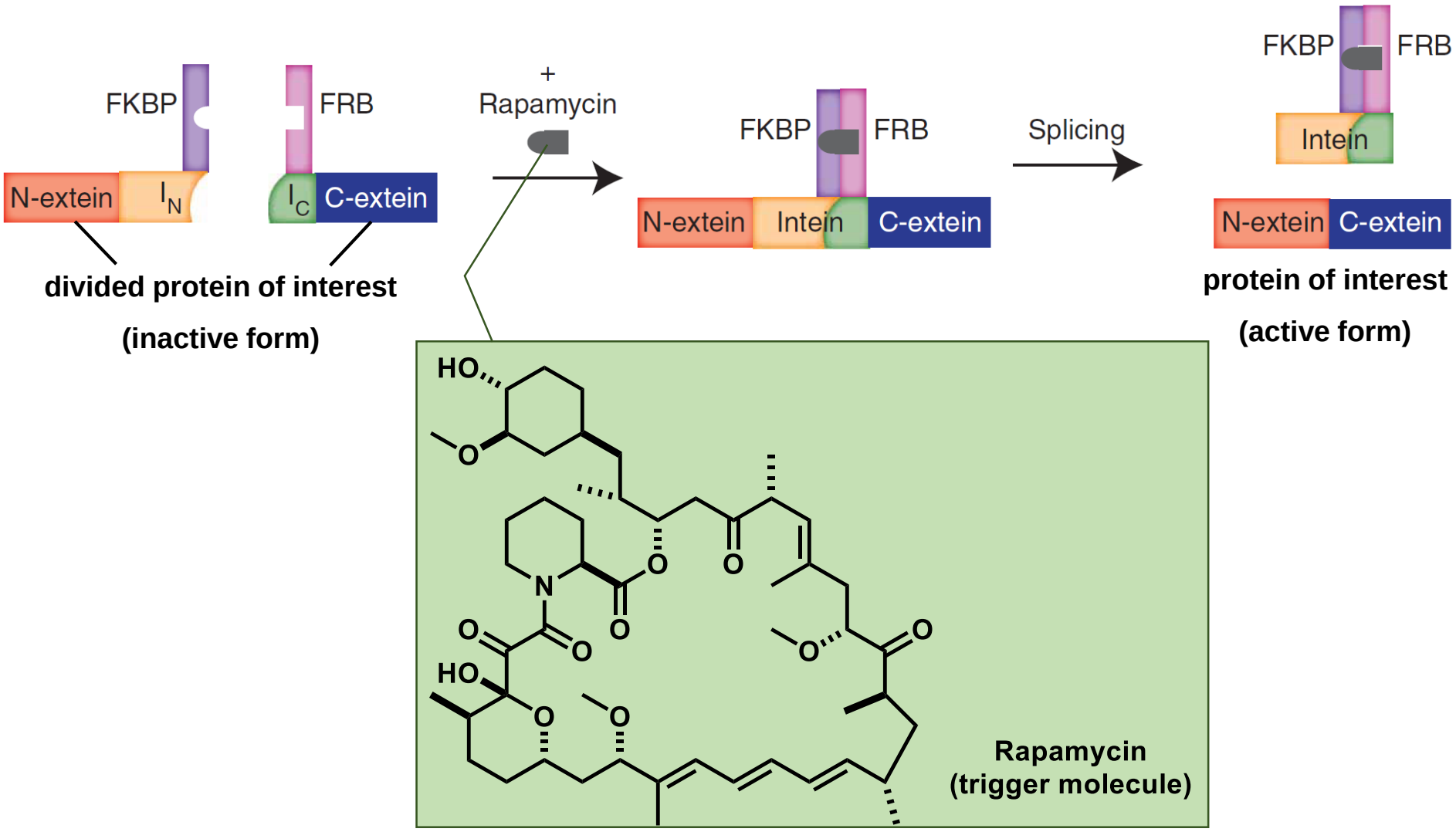


Seyedsayamdost, M. R.; Yee, C. S. and Stubbe, J. *Nat. Protocol* **2007**, 2, 1225

Lu, W *et al. Mol. Cell* **2001**, 8, 759

Conditional Protein Splicing

Intein: Intron in protein processing/ **FKBP**: FK506-binding protein 12/ **ERB**: FKBP-rapamycin binding protein

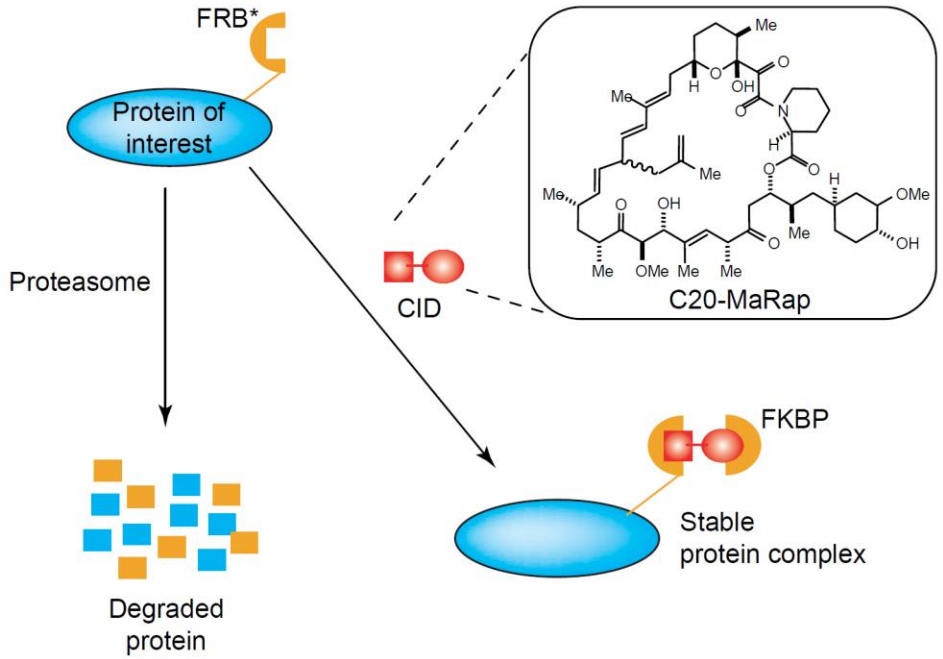


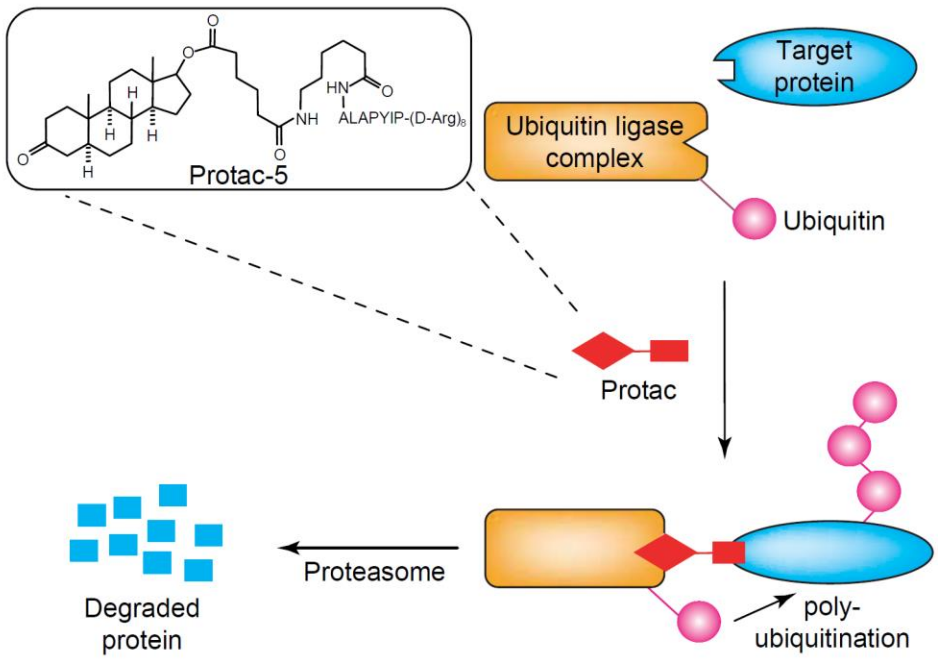
Mootz, H. D.; Blum, E. S.; Tyszkiewicz, A. B. and Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125*,10561.
Schwartz, E. C.; Saez, L.; Young, M. W. and Muir, T. W. *Nat. Chem. Biol.* **2007**, *3*, 50.

Conditional Protein Degradation

(a) Chemical rescue

(b) Chemical induction





FRB*: a mutant FRB domain that is constitutively degraded by the proteasome but can be stabilized by rapamycin or close analogs

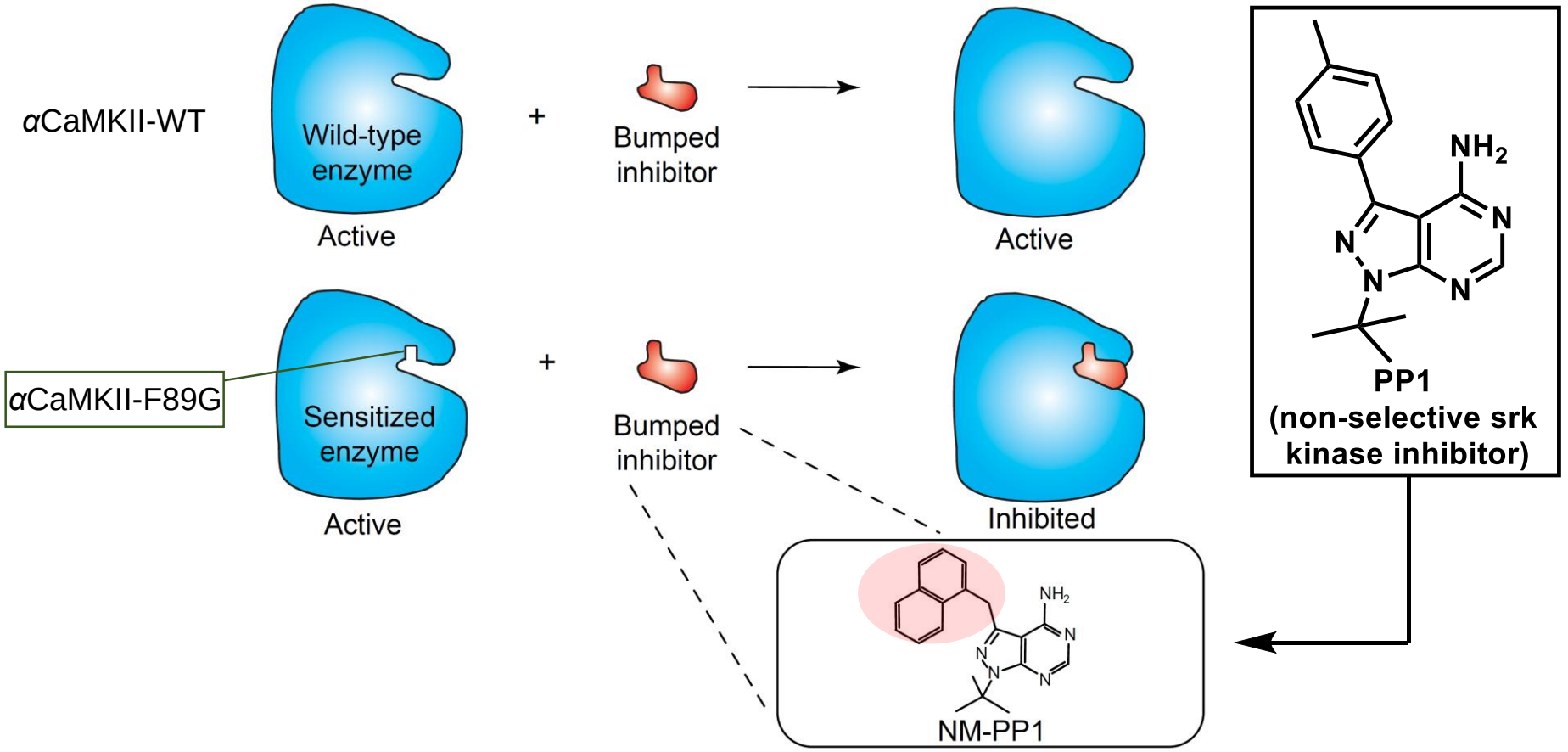
Protac: Proteolysis-targeting Chimera = bifunctional small molecule for removal of target protein from the cell

CID = Chemical Inducer of Dimerization

- **Protac-5**: a dihydrotestosterone- ALAPYIP (E3 recognition domain)-polyarginine conjugate
- **C20-MaRap**: a rapamycin derivative

Hahn, M.E.; Muir, T. W. *TRENDS in Biochem. Sci.* **2005**, 30, 1
Stankunas, K.; Bayle, J. H.; Gestwicki, J. E.; Lin, Y-M.; Wandless, T. J.; Crabtree, G. R. *Mol. Cell* **2003**, *12*, 1615
Schneekloth, J. S. Jr.; Fonseca, F. N.; Koldobskiy, M.; Mandal, A.; Deshaies, R.; Sakamoto, K. and Crews, C. M. *J. Am. Chem. Soc.* **2004**, *126*, 3748

“Bump-and-Hole” Strategy



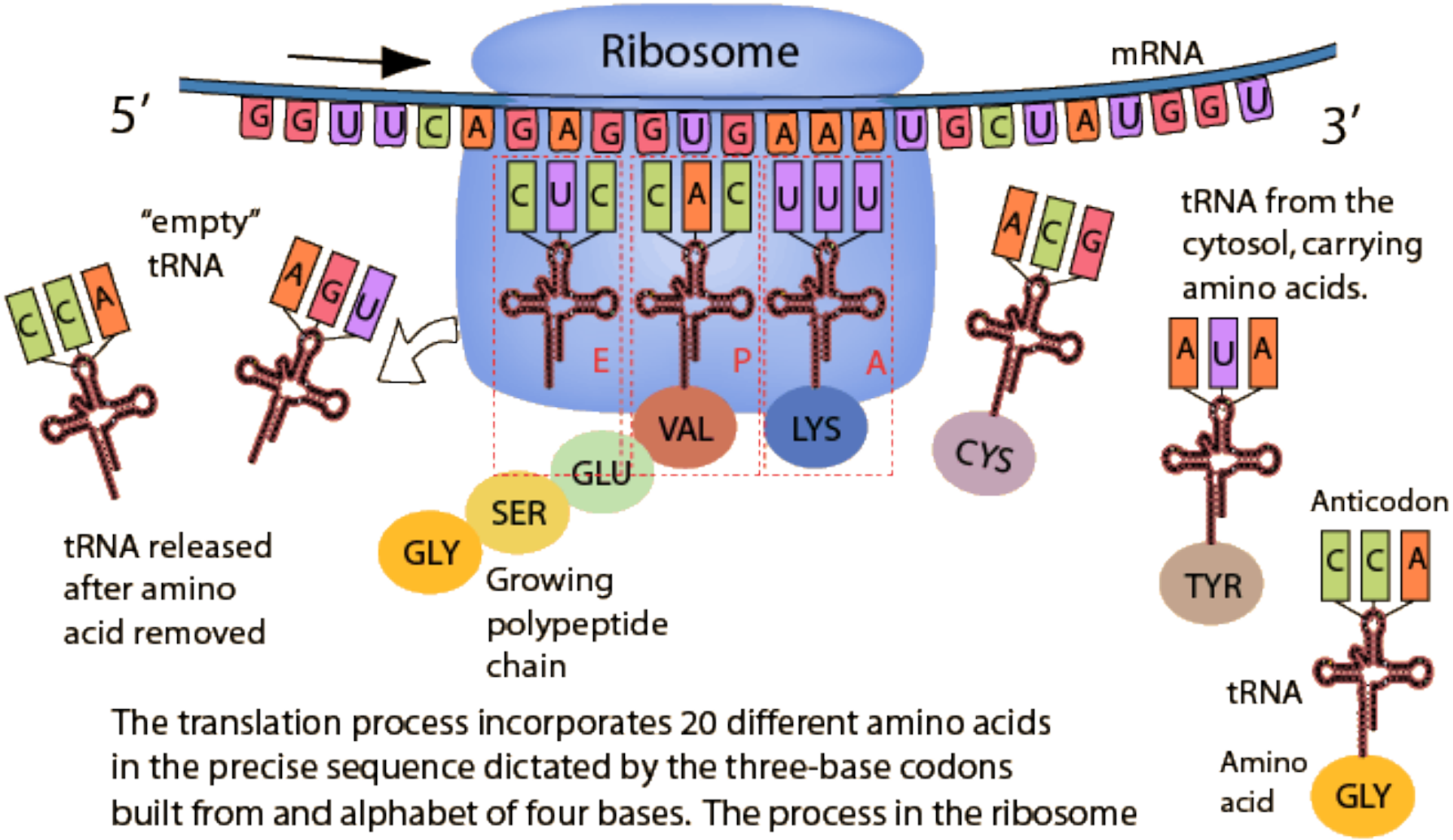
“Bumped” inhibitor (NM-PP1) × genetically “Holed” enzyme enabled highly selective inhibition

Hahn, M.E.; Muir, T. W. *TRENDS in Biochem. Sci.* **2005**, 30, 1
Wang. H *et al. Proc. Natl. Acad. Sci. U. S. A.* **2003**,*100*, 4287
Belshaw, P. J.; Schoepfer, J. G.; Liu, K-Q.; Morrison, K. M. and Schreiber, S. L. *Angew. Chem. Int. Ed.* **1995**, *34*, 2129

Outline

1. Introduction
2. Incorporation of unnatural amino acids
3. Bioorthogonal ligand tethering (BOLT)
4. Protein-control using Staudinger reduction (Main Paper)
5. Summary

Translation (Fundamental)



The translation process incorporates 20 different amino acids in the precise sequence dictated by the three-base codons built from and alphabet of four bases. The process in the ribosome builds the polypeptide chains tha will become proteins.

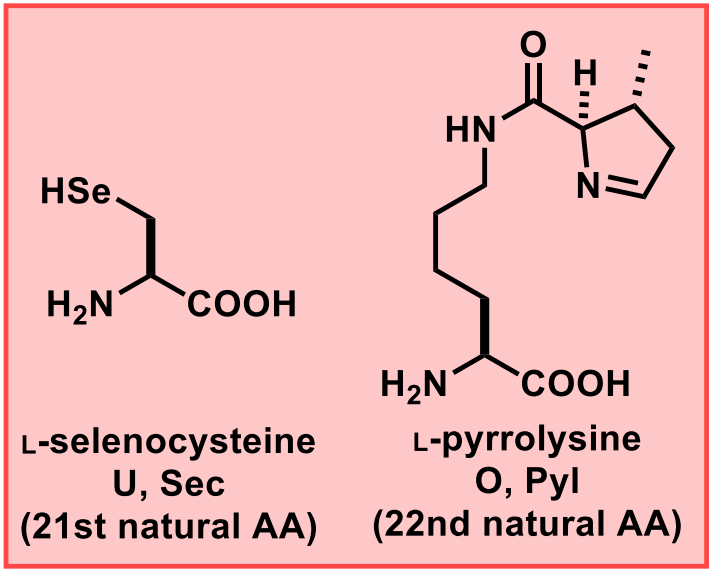
Aminoacyl tRNA is synthesized by a corresponding aminoacyl tRNA synthetase (example: Arginine-tRNA/arginyl-tRNA synthetas (ArgRS))

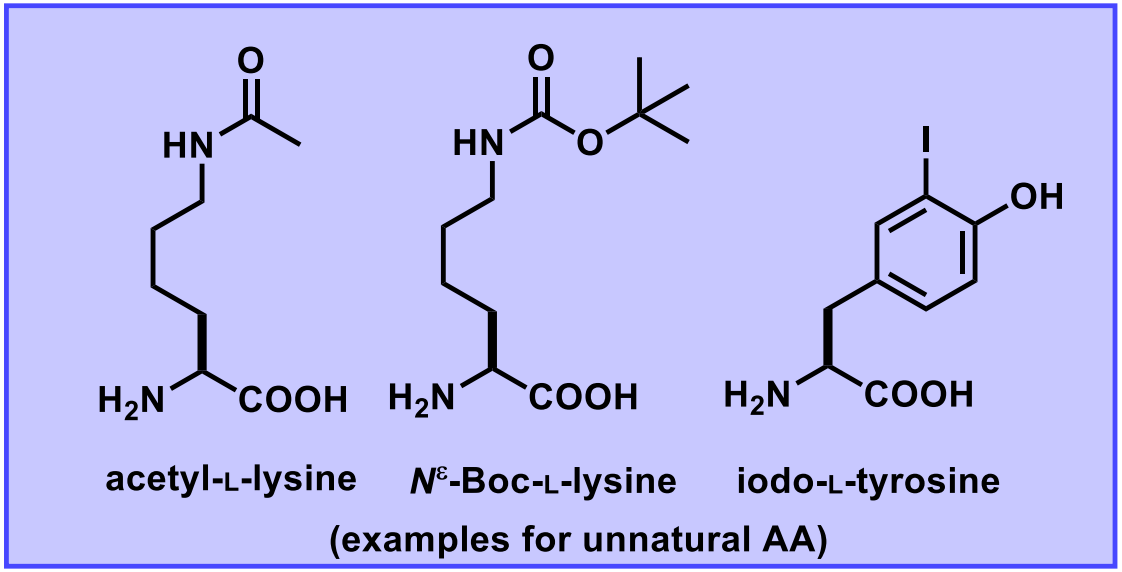
http://hyperphysics.phy-astr.gsu.edu/hbase/organic/translation.html

Alloproteins & Unnatural Amino Acids

- Alloproteins: proteins substituted with unnatural amino acids
- Unnatural amino acids: amino acids not naturally encoded or found in the genetic code of any organisms

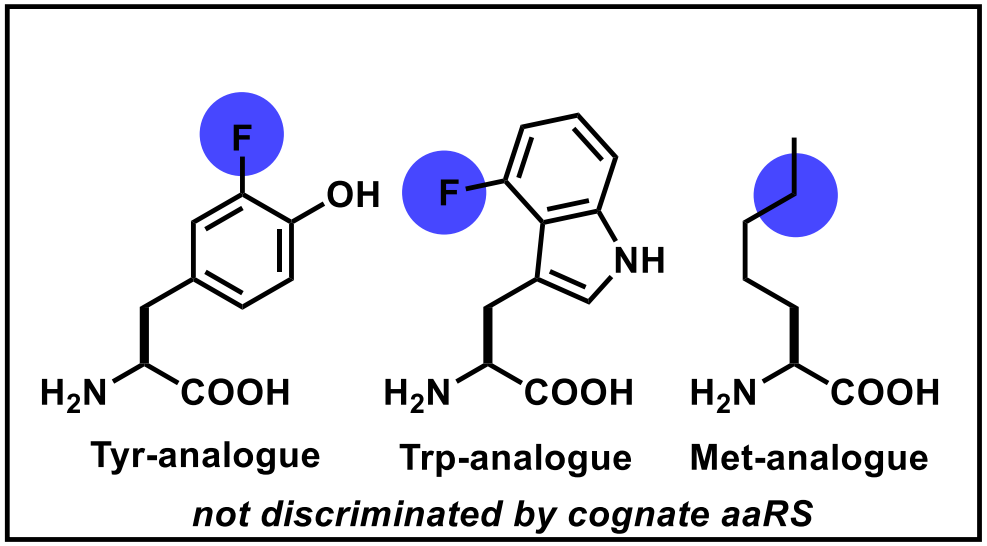
(not including *selenocysteine* and *pyrrolysine*)

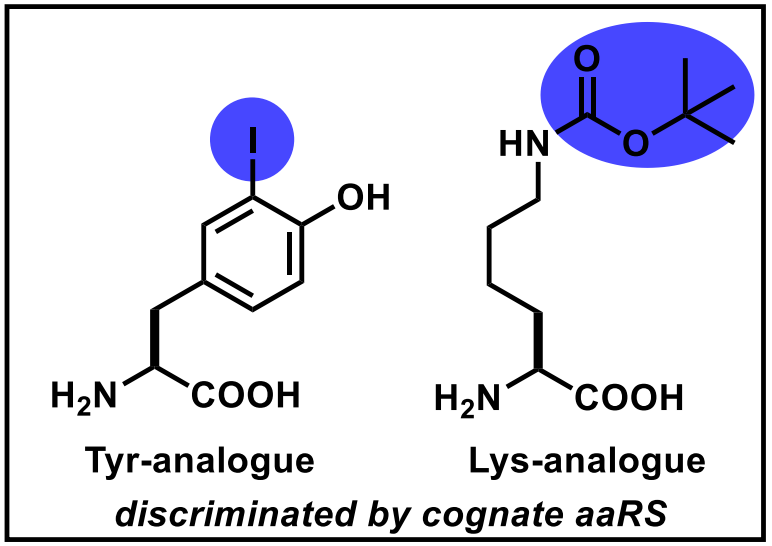




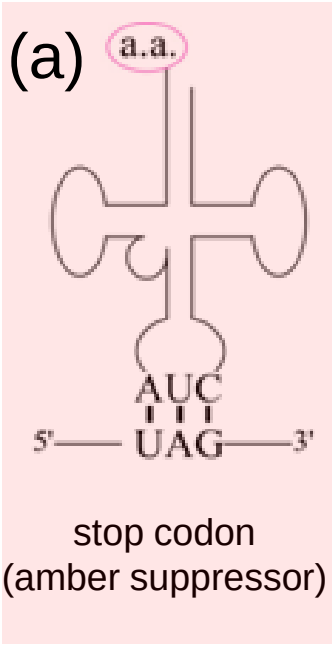
Koide, H. *et al. Proc. Natl. Acad. Sci. U. S. A.* **1998**, *85*, 6237
Sakamoto, K. *et al. Nucleic Acids Res.* **2002**, 30, 4692

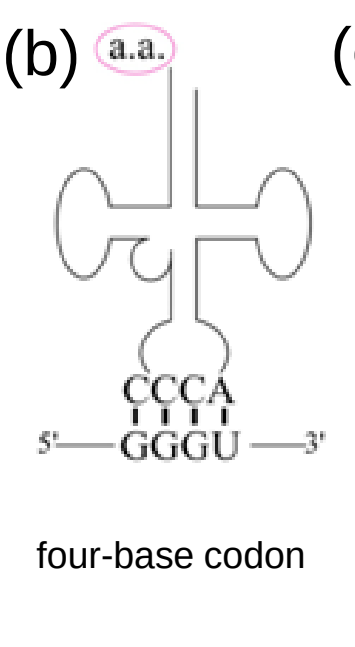
Expansion of the Genetic Code System

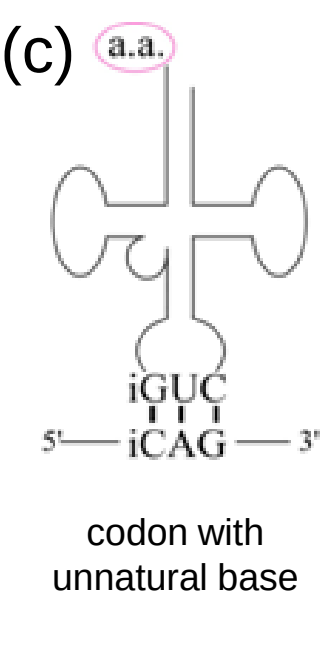




	T		C		A		G	
T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
	TTC		TCC		TAC		TGC	
	TTA	Leu	TCA		TAA	STOP	TGA	STOP
	TTG		TCG		TAG	STOP	TGG	Trp
C	CTT	Leu	CCT	Pro	CAT	His	CGT	Arg
	CTC		CCC		CAC		CGC	
	CTA		CCA		CAA	Gln	CGA	
	CTG		CCG		CAG		CGG	
A	ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser
	ATC		ACC		AAC		AGC	
	ATA		ACA		AAA	Lys	AGA	Arg
	ATG	Met	ACG		AAG		AGG	
G	GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly
	GTC		GCC		GAC		GGC	
	GTA		GCA		GAA	Glu	GGA	
	GTG		GCG		GAG		GGG	

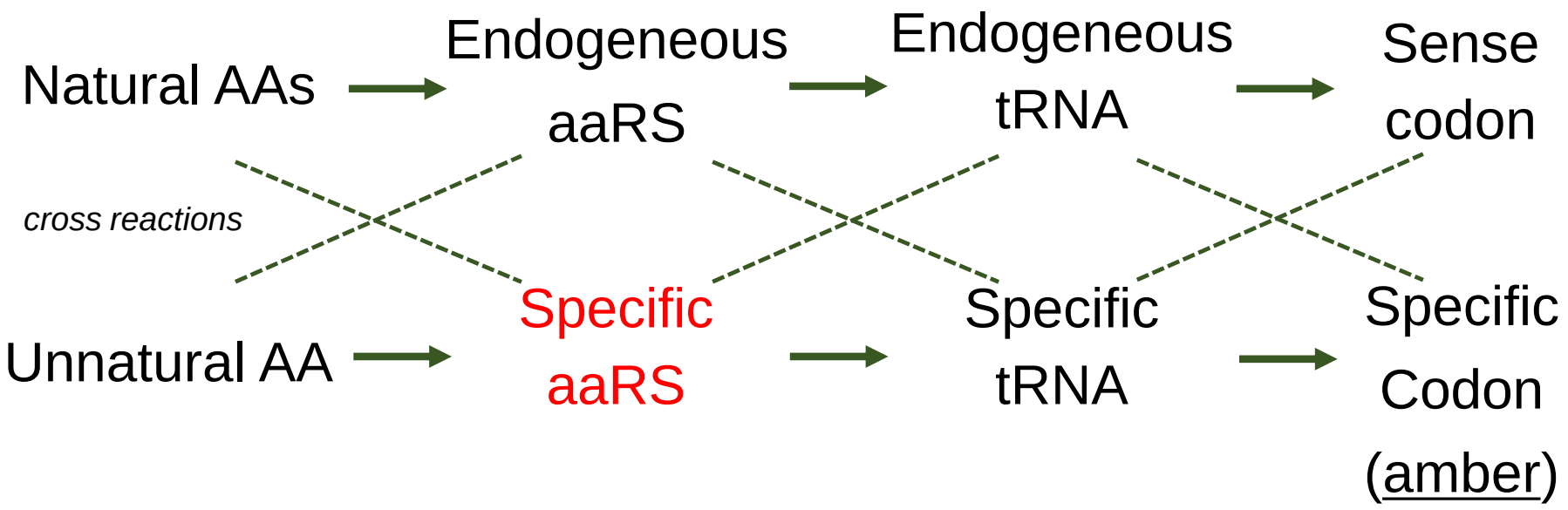






Koide, H. *et al. Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 6237.

Expansion of the Genetic Code System



To make an “orthogonal” tRNA/aaRS pair, aaRSs from bacteria and archaea are useful

1. Tyrosyl-tRNA synthase (TyrRS) from bacteria (*E. coli etc.*)

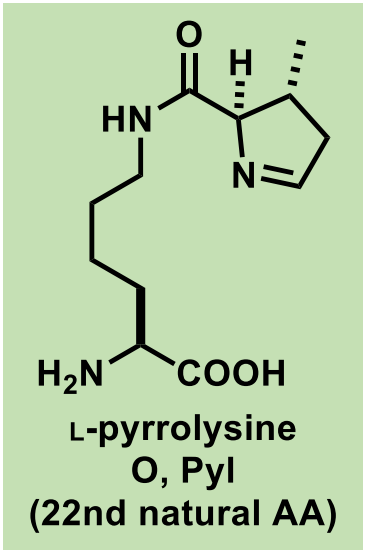
Incorporation of unnatural Tyr-derivatives

2. Pyrrolysyl-tRNA synthase (PylRS) from archaea (*Methanosarcina barkeri*)

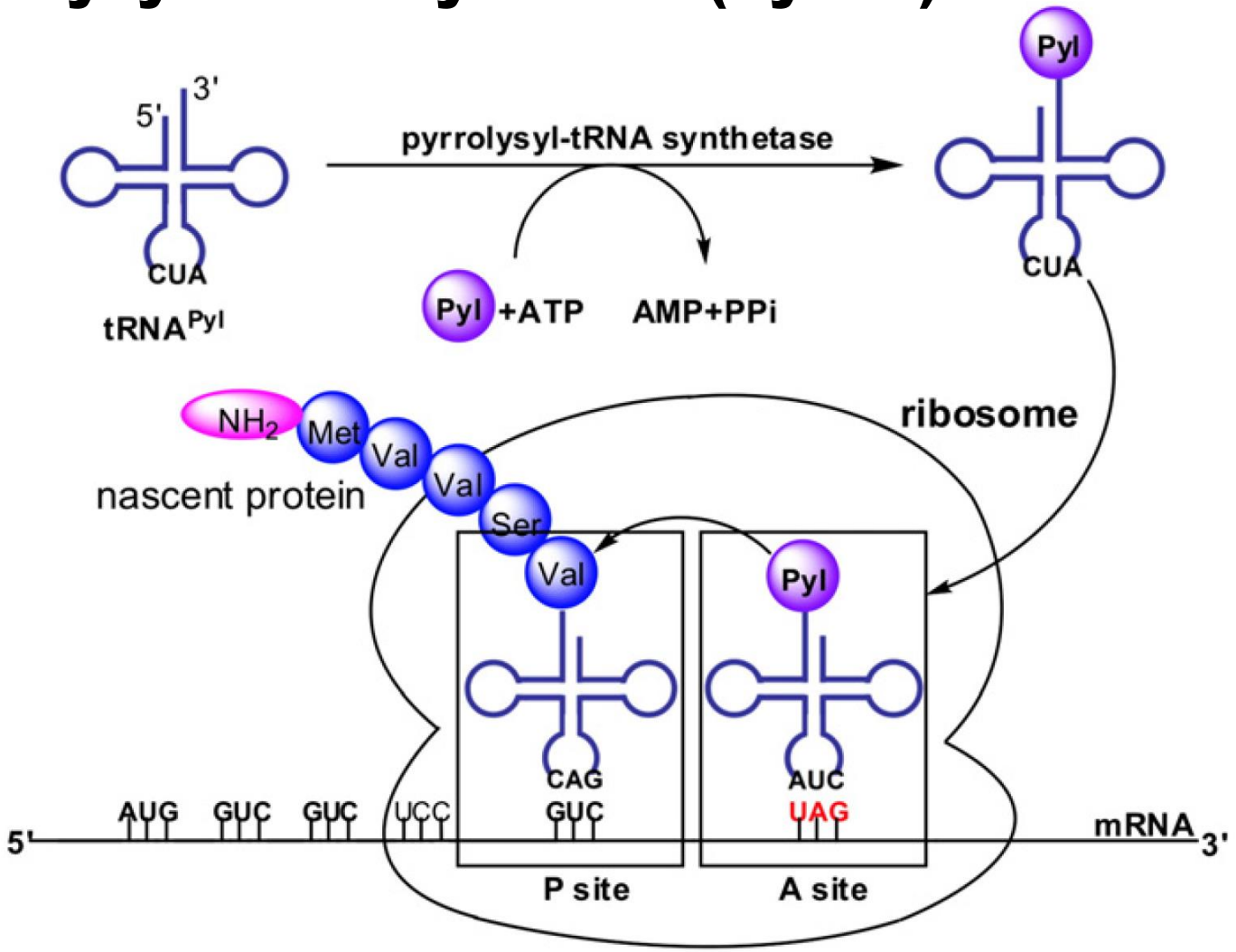
Incorporation of unnatural Lys-derivatives

Sakamoto, K *et al. Nucleic Acids Res.* **2002**, *30*, 4692
Yanagisawa, T. *et al. Chem. Biol.* **2008**, *15*, 1187

Pyrrolysyl tRNA Synthase (PyrRS)



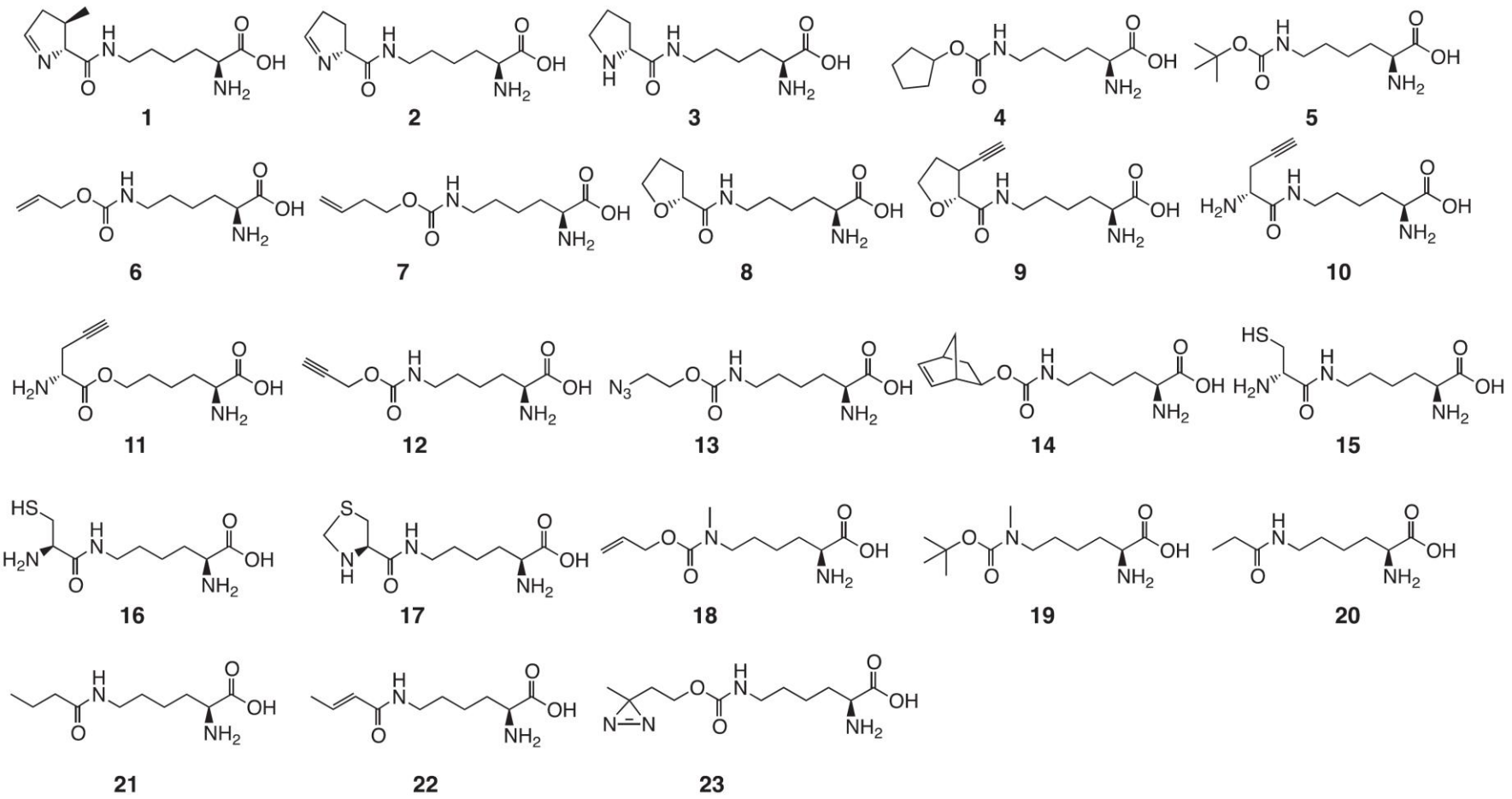
L-pyrrolysine
O, Pyl
(22nd natural AA)



- PylRS is a protein encoded only in DNA of Methanogens
- $tRNA^{Pyr}$ has an anti-codon “CUA” complementary to amber codon (UAG)
- $tRNA^{Pyr}$ is not recognized by *E. Coli* and mammalian cell’s aaRSs

Yanagisawa, T. *et al. Chem. Biol.* **2008**, *15*, 1187
Wan, W.; Tharp, J. M.; Liu, W. R. *Biochim Biophys Acta.* **2014**,*1844*, 1059

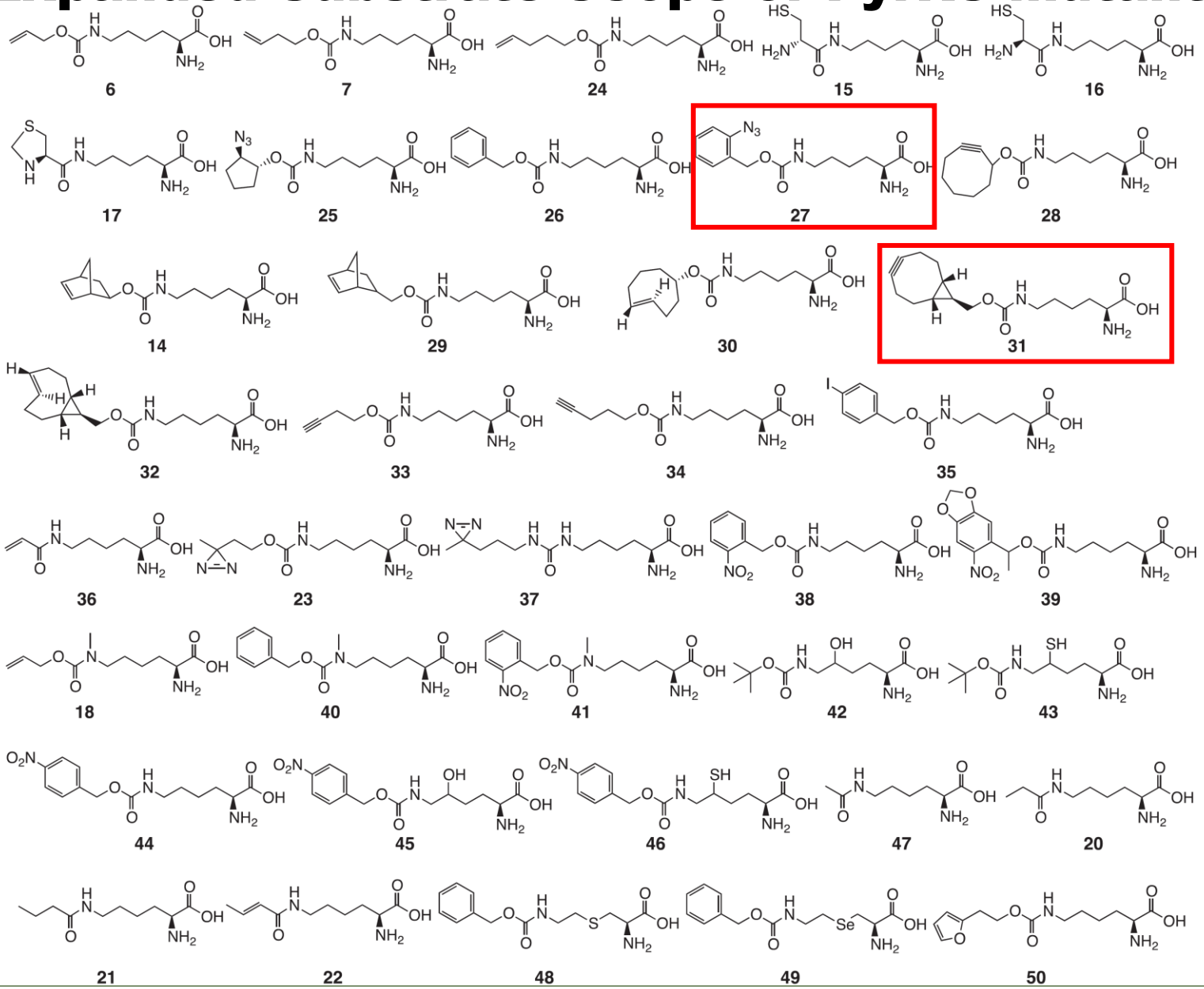
Substrate Scope of the native PyrRS



To expand the range of applicable non-natural amino acids, some active mutants have been selected from the randomly mutated library
(Also see *appendix-1* for Phe derivatives)

Wan, W.; Tharp, J. M.; Liu, W. R. *Biochim Biophys Acta.* **2014**,*1844*, 1059

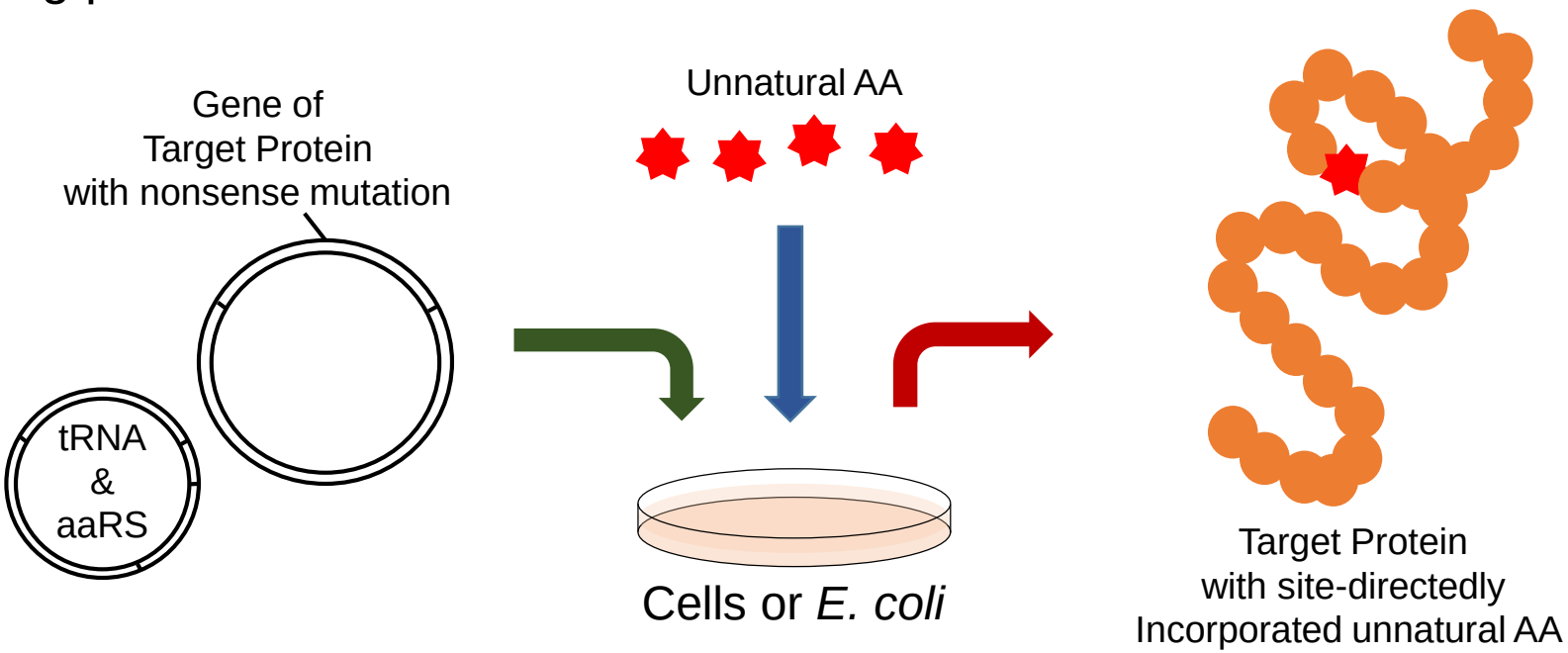
Expanded Substrate Scope of PyrRS mutants



Wan, W.; Tharp, J. M.; Liu, W. R. *Biochim Biophys Acta.* **2014**,*1844*, 1059

How to Use the “Orthogonal” tRNA/aaRS Pair

1. Preparing the plasmid DNA of a target protein with required nonsense codons
2. Adding plasmid DNAs and the unnatural amino acid



Applications:

- X-ray analysis (iodo-Tyr)
- Selective functionalization (biotin, fluorophore etc)
- Protein-protein interaction analysis with a photo-cross linker
- Post-translational modification analysis
- *In cell* protein-manipulation

Outline

1. Introduction
2. Incorporation of unnatural amino acids
3. Bioorthogonal ligand tethering (BOLT)

ARTICLES
PUBLISHED ONLINE: 18 MAY 2015 | DOI: 10.1038/NCHEM.2253
nature chemistry

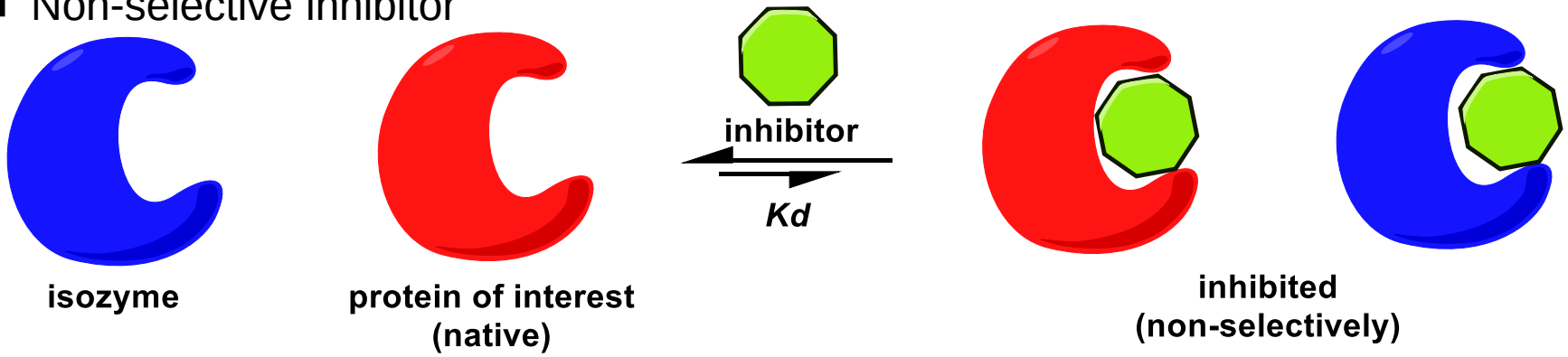
Selective, rapid and optically switchable regulation of protein function in live mammalian cells

Tsai, Y-H.; Essig, S.; James, J. R.; Lang, K.; Chin, J. W. *Nat. Chem.* **2015**, *7*, 554-561

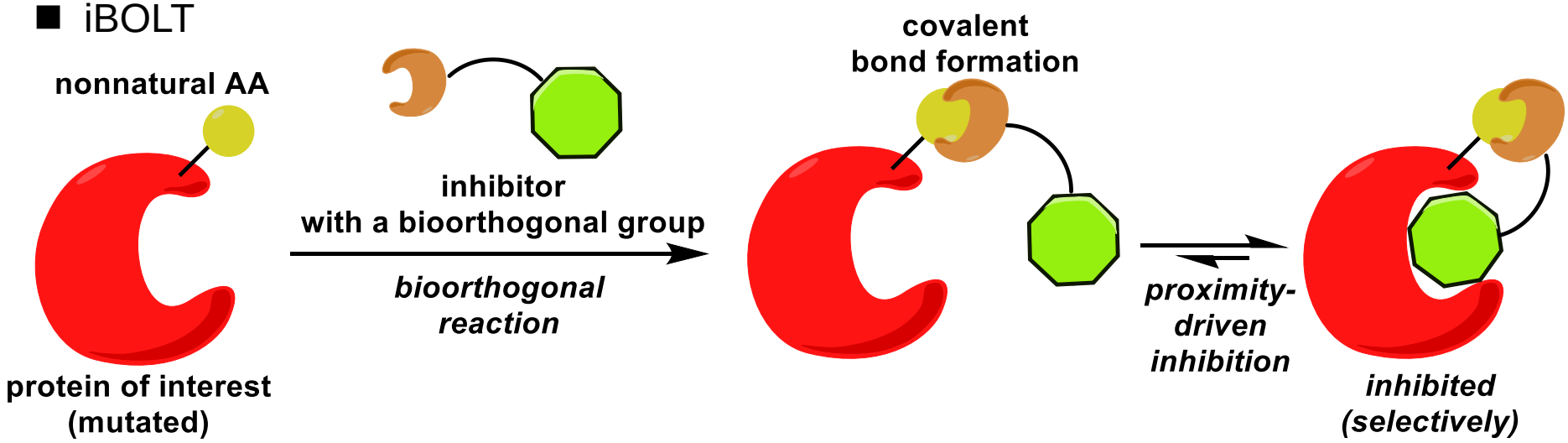
4. Protein-control using Staudinger reduction (Main Paper)
5. Summary

Inhibition via Bioorthogonal Ligand Tethering (iBOLT)

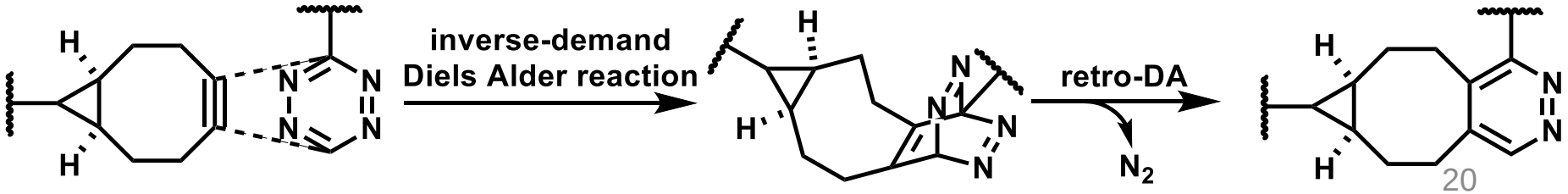
- Non-selective inhibitor



- iBOLT



- *This approach using "tetrazine-ligation" was applied to MEK kinase isozymes*



MEK Kinases and Design of an Inhibitor Conjugate

- MEK1 and MEK2:

-key kinases in the MAP kinase signaling pathway

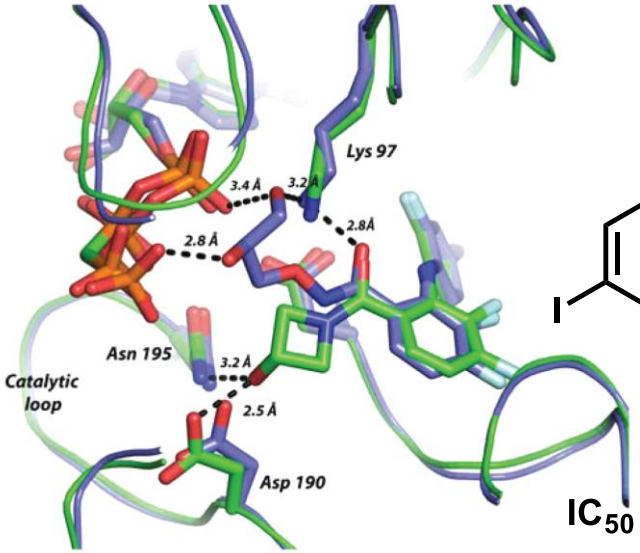
-they share 82% sequence identity

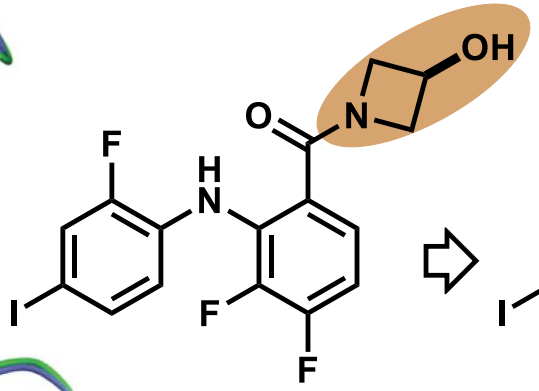
-selective inhibition of MEK1 or MEK2 has not been achieved

-knockout or knockdown approaches were not applicable

-"bump and hole" strategy was failed

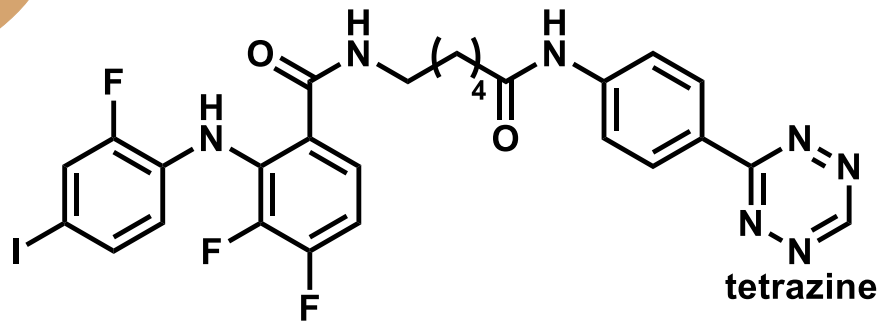
- Design of an Inhibitor-triazine Conjugate





inhibitor 2

IC_{50} = 44 nM against MEK1

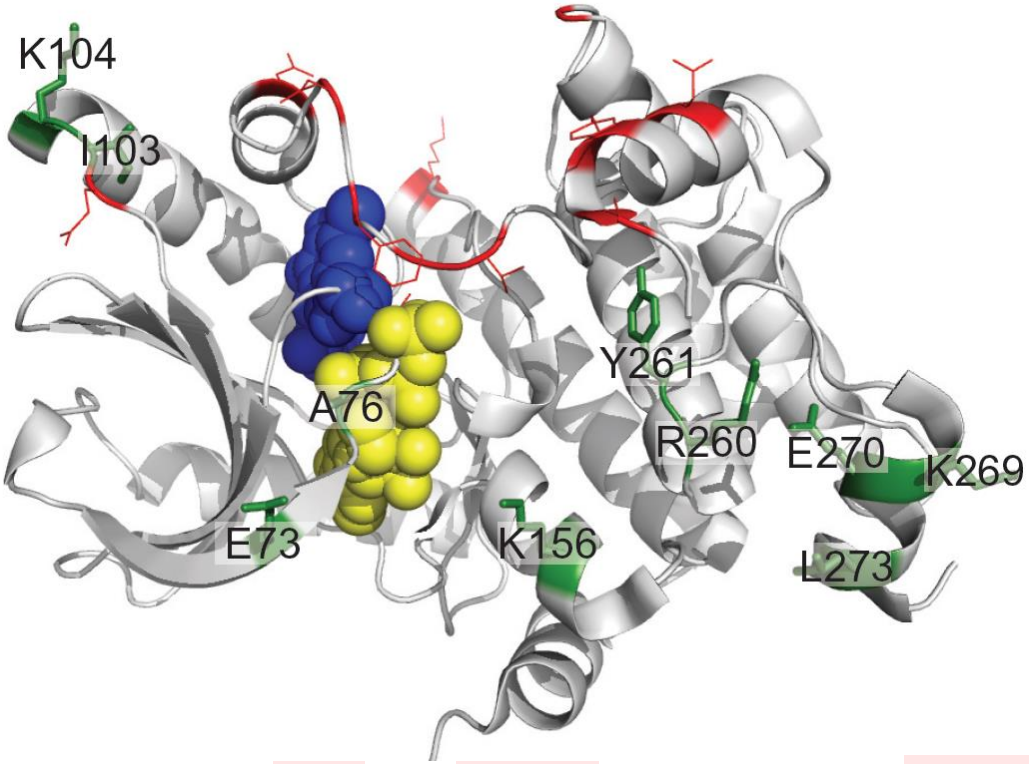


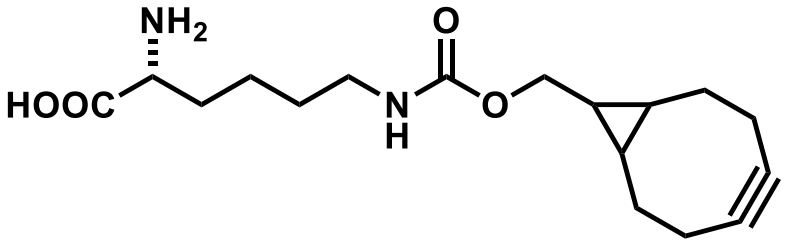
derivative 3

(IC_{50} *in cell* against MEK1 > 10 µM)

Rice, K. D. *et al. ACS Med Chem Lett* **2012**, *3*, 416

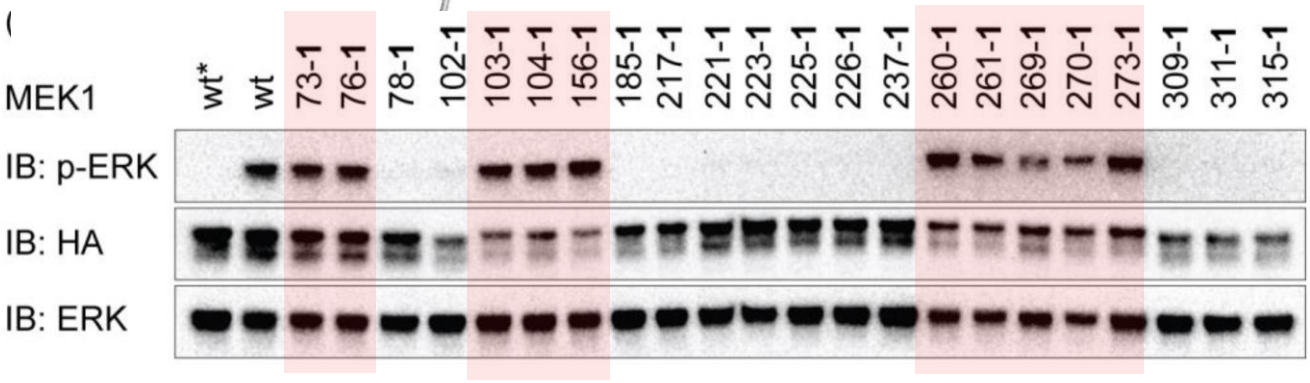
Identifying Active MEK1 Variants with a Non-natural Amino Acid





unnatural AA (1)

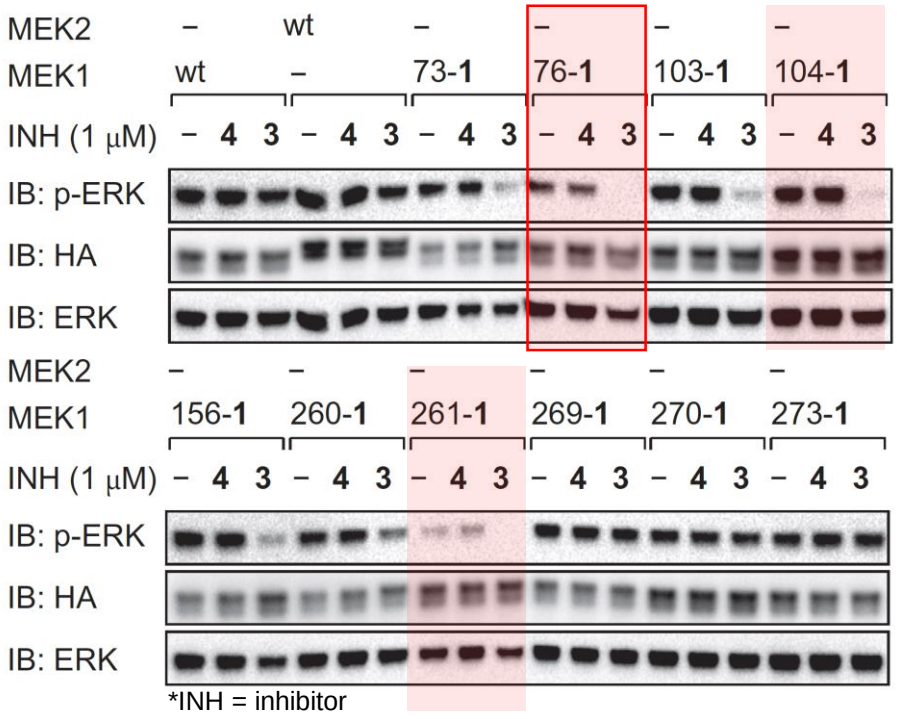
*ATP and inhibitor **2** are shown as yellow and blue spheres, respectively.
**Twenty-two of solvent-exposed AA residues were identified within 40 Å from the amide in the inhibitor.



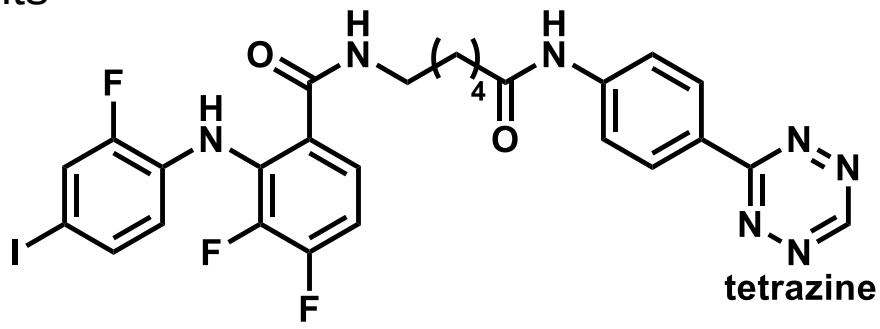
<u>E73, A76, I103, K104, K156, R260, Y261, K269, E270 and I 273 were permissive</u>

Investigation into iBOLT of MEK1 Variants

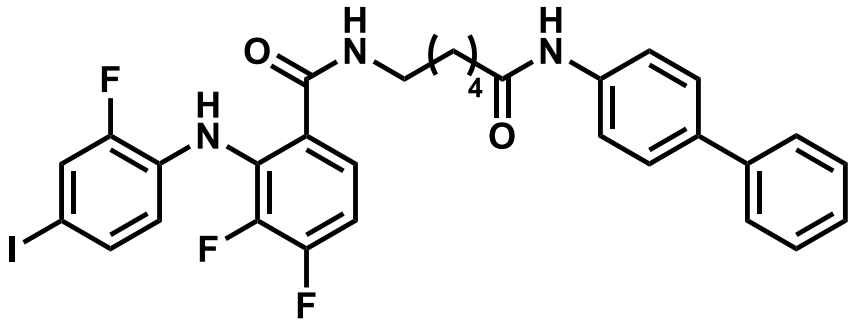
- Derivative **3** induced inhibition of MEK variants



*INH = inhibitor

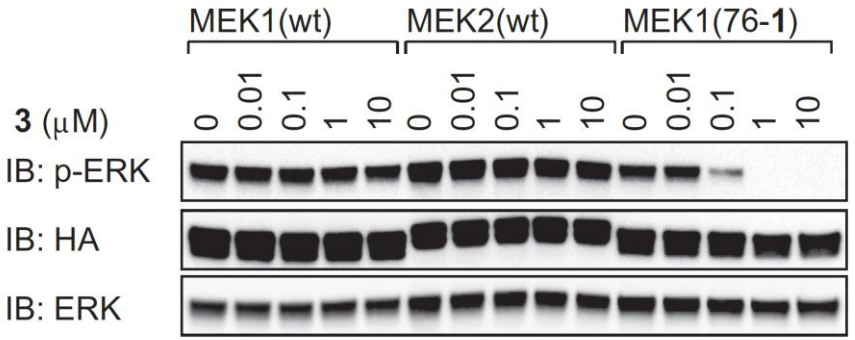


derivative 3

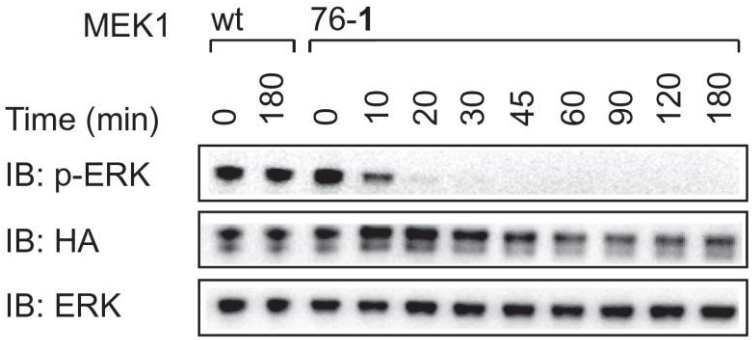


derivative 4

- Concentration dependence (76-1)



- Time Dependence (76-1)

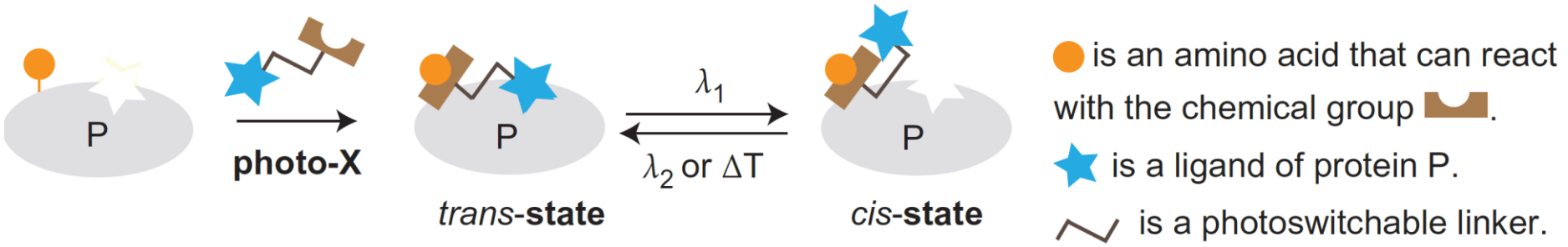


Optical Switching of MEK1 with “Photo-BOLT”

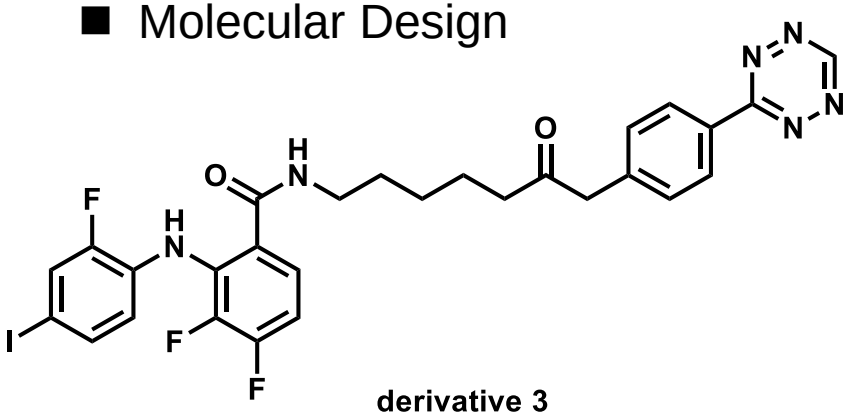
Inhibition of MEK1 by iBOLT was sensitive to (1) position of tethering and (2) linker length

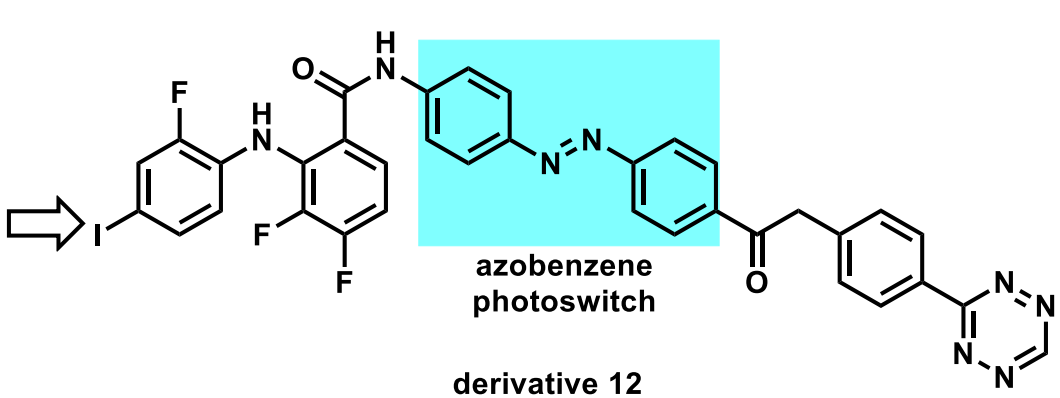
A *photoisomerizable linker* enables the reversible switching of MEK1 function

- Concept



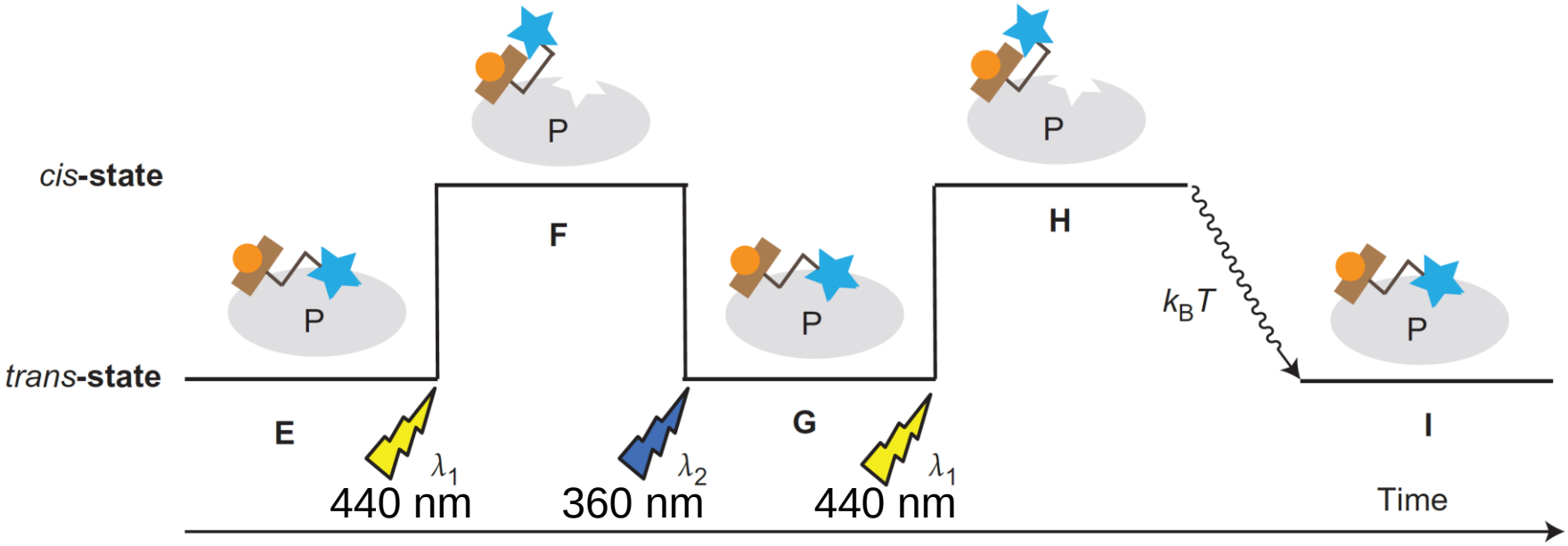
- Molecular Design

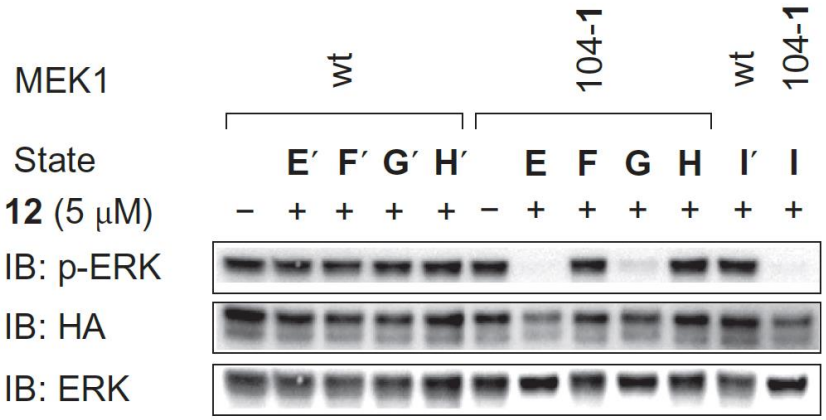


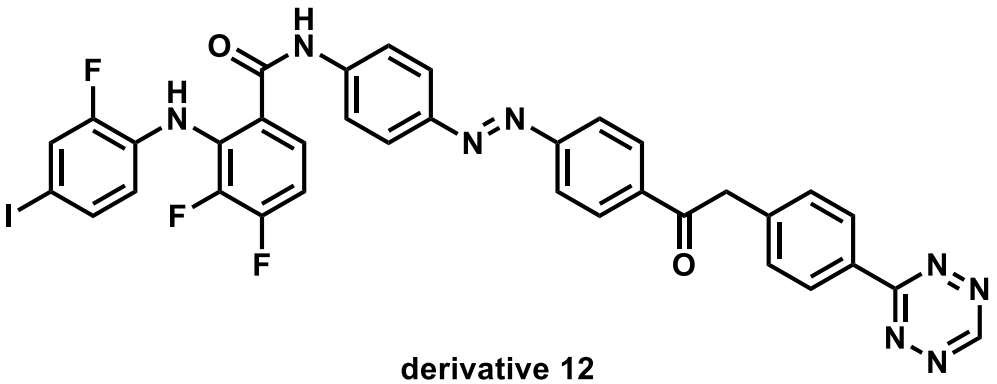


Optical Switching of MEK1 with “Photo-BOLT”

- MEK mutants screening showed that the 104-1 mutant is appropriate for derivative **12**







derivative 12

Outline

1. Introduction
2. Incorporation of unnatural amino acids
3. Bioorthogonal ligand tethering (BOLT)
4. Protein-control using Staudinger reduction (Main Paper)

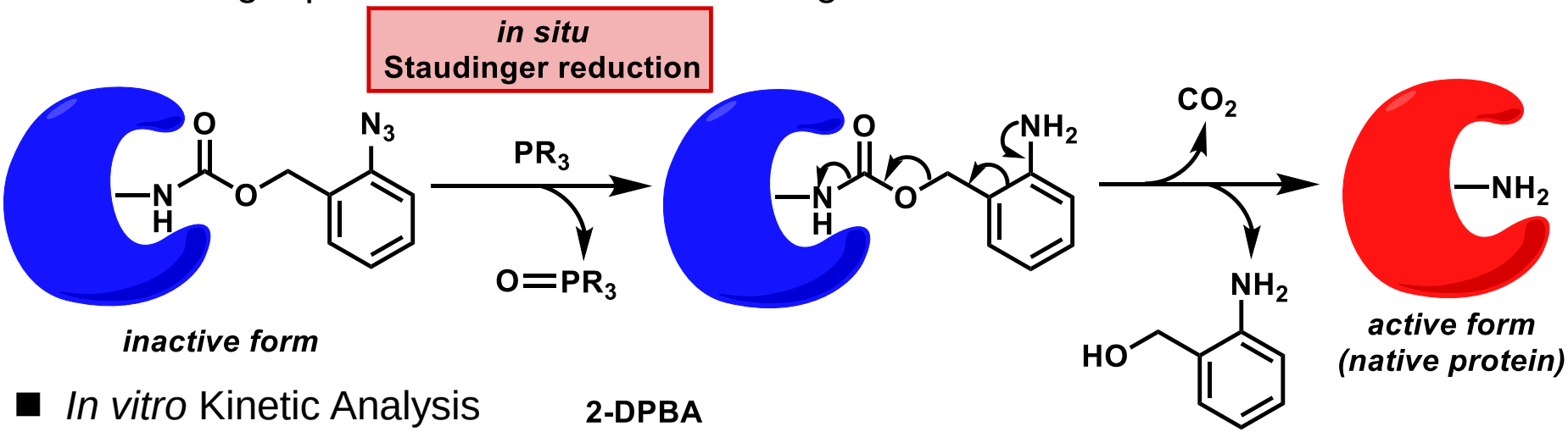
Small-molecule control of protein function through Staudinger reduction

Luo, J.; Liu, Q.; Morihiro, K. and Deiters, A. *Nat. Chem.* **2016**, online publication

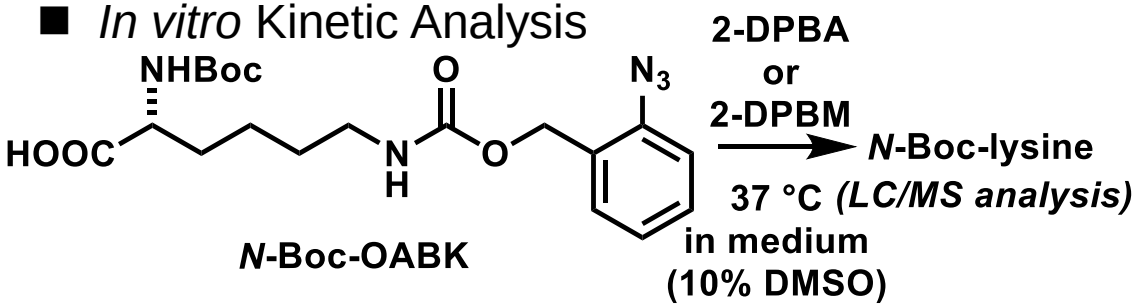
5. Summary

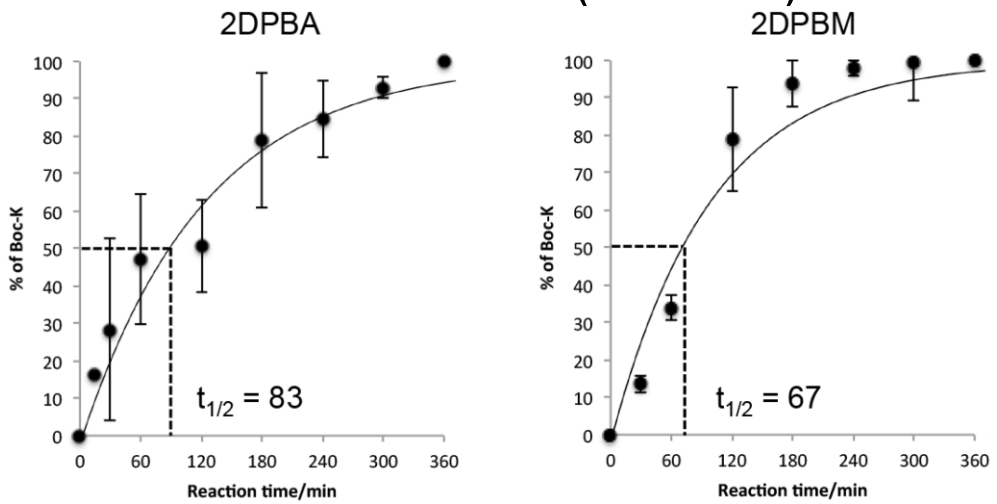
Staudinger Reduction of "OABK"

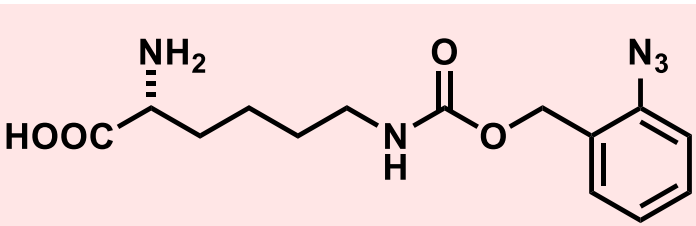
- Controlling a protein function with Staudinger reduction

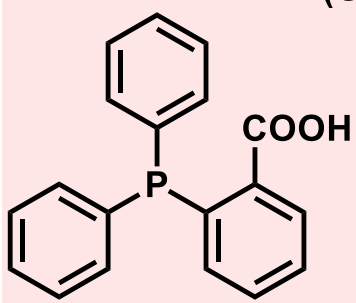


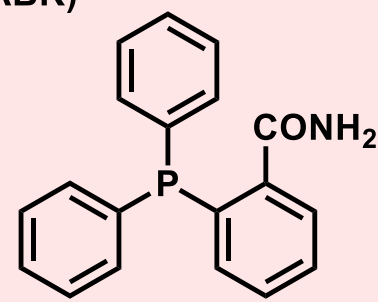
- *In vitro* Kinetic Analysis





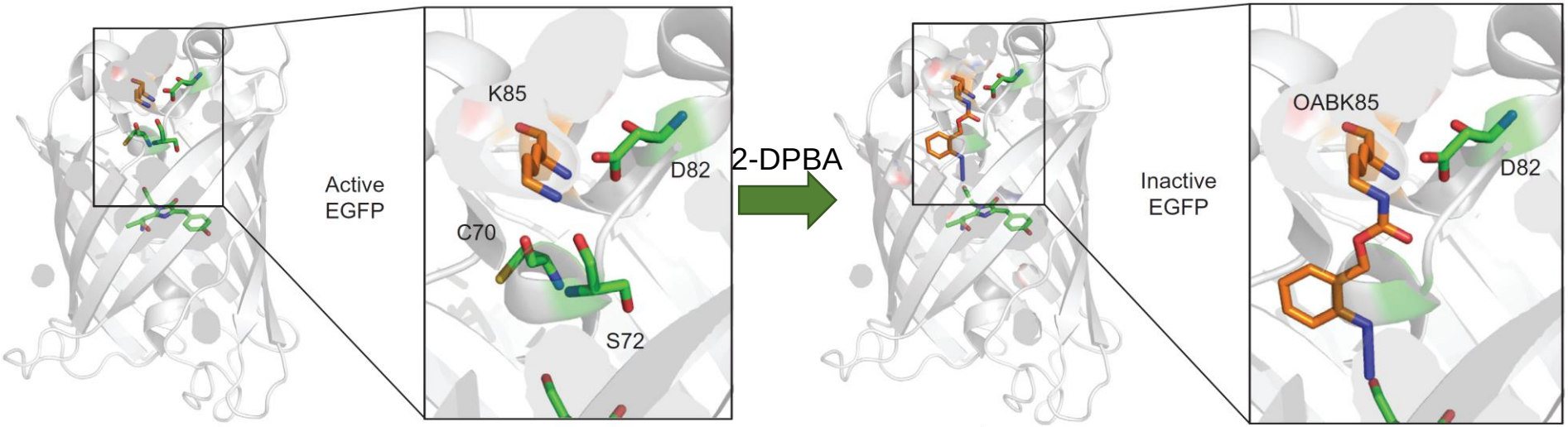




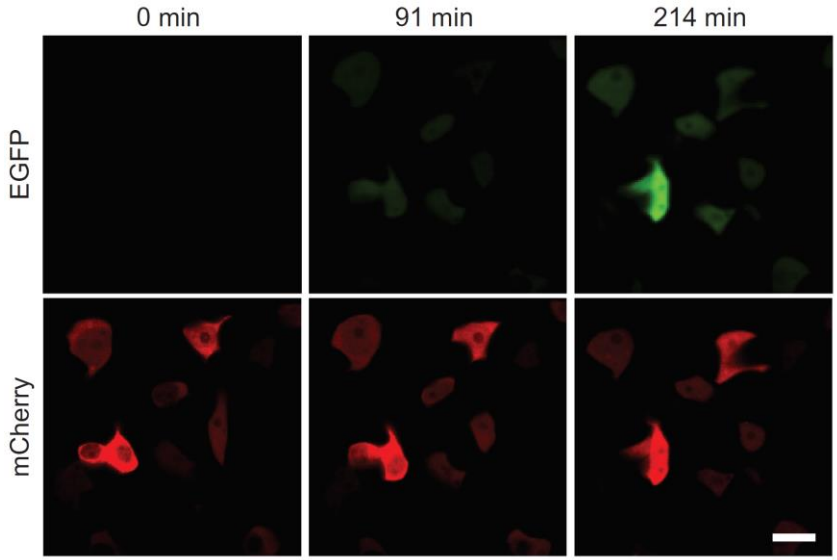


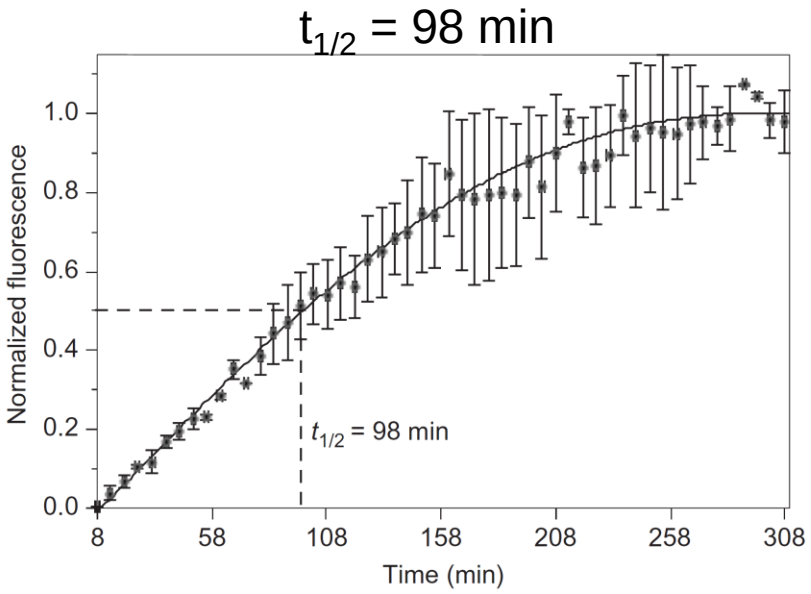
Fluorescent Protein (eGFP) Activation

- Design of the mutation – K85 is the crucial residue (hydrogen-bond formation)



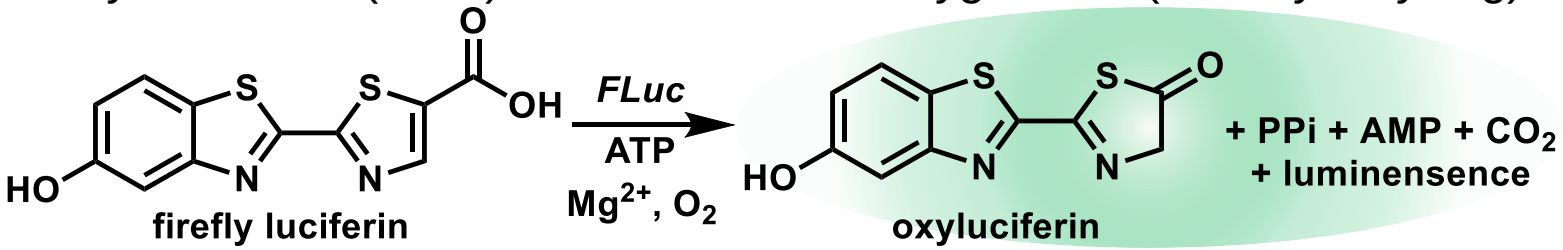
- Results



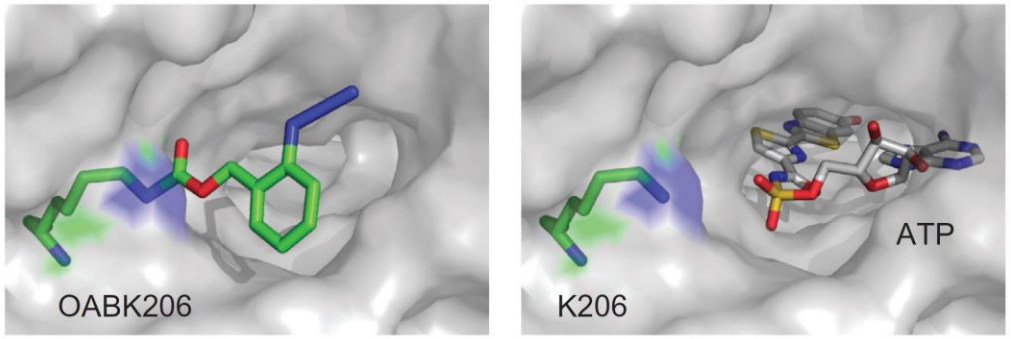


Enzyme (Luciferase) Activation and *in cell* Phosphine Screening

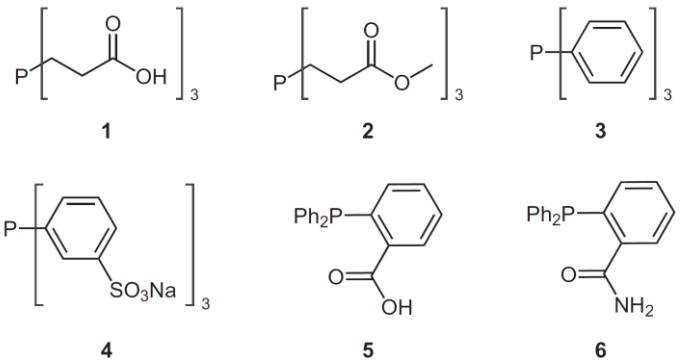
- Firefly luciferase (FLuc) - luciferin 4-monooxygenase (ATP-hydrolysing)



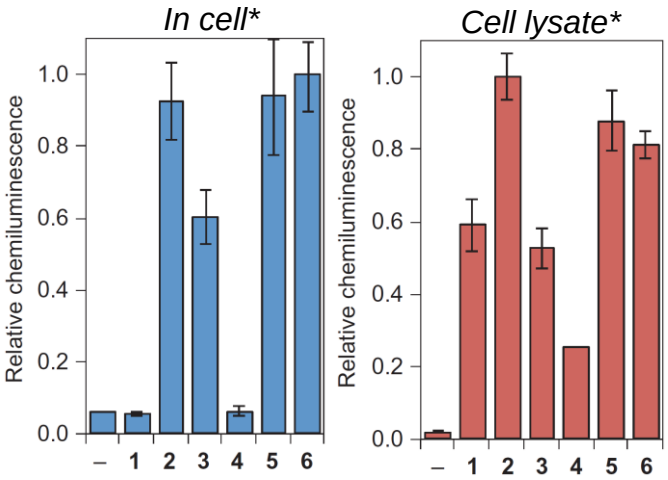
- Design of Mutation (blocking the ATP pocket)

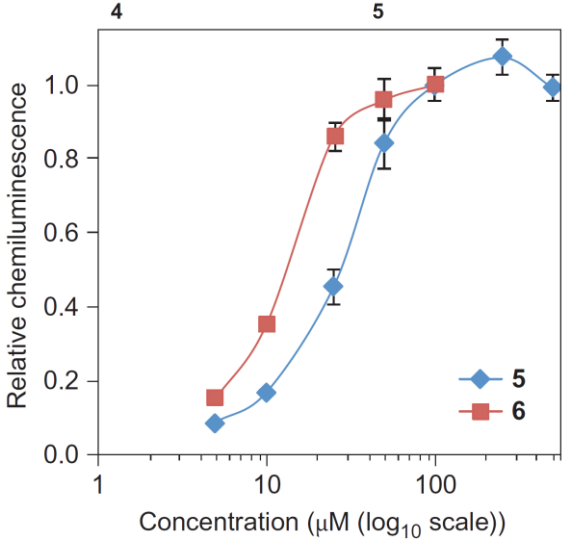


- Phosphine Screening



- Results



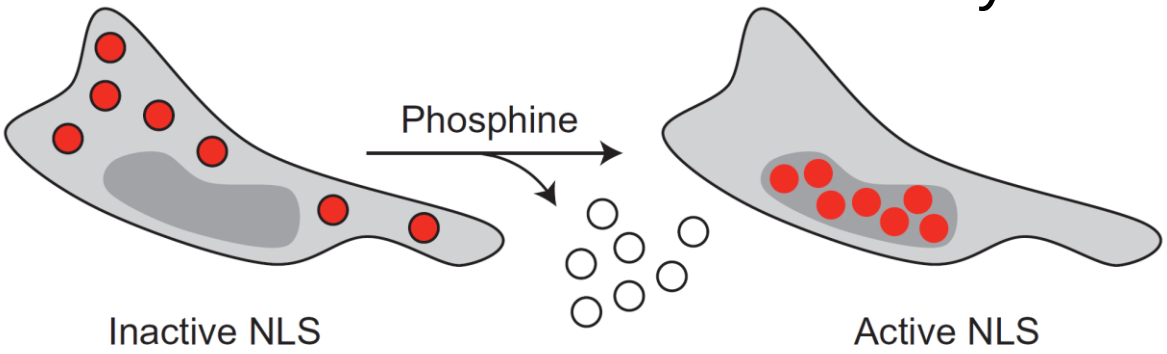


*Phosphines were used in maximum concentration based on the solubility and cell viability

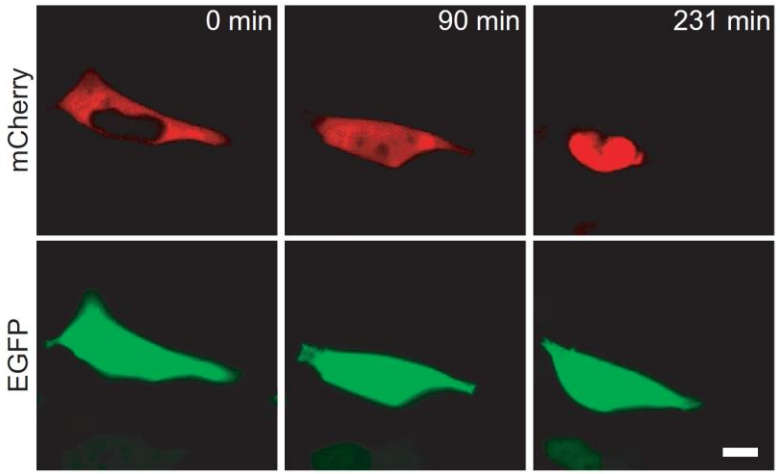
Protein Translocation of SATB1

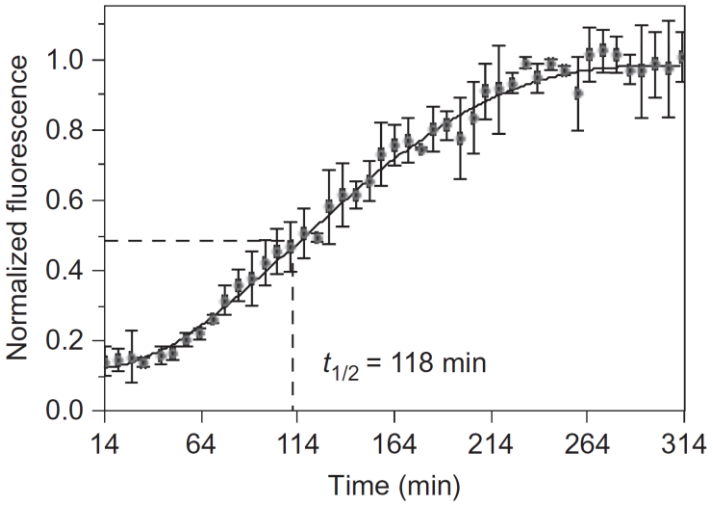
- SATB1 – A protein involving chromatin-loop remodeling
- NLS (Nuclear Localization Signal) – NLS sequence of SATB1 is residues 20-40
- Design of Mutation

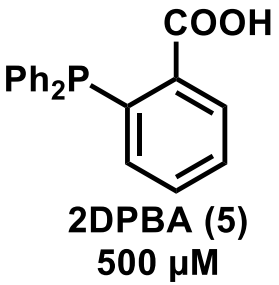
SATB1-K29OABK-mCherry



- Results

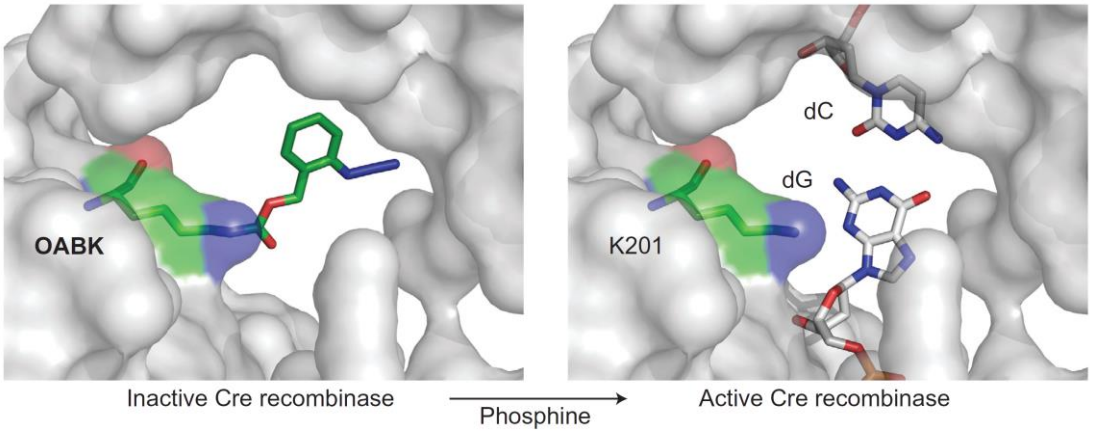


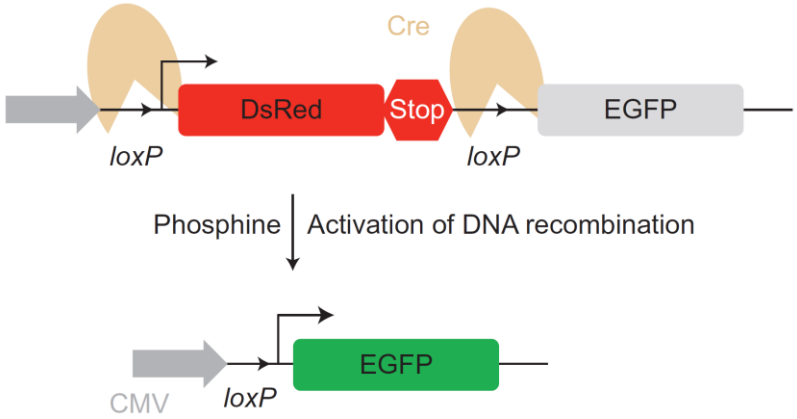




DNA Recombination Using the Cre-*loxP* System

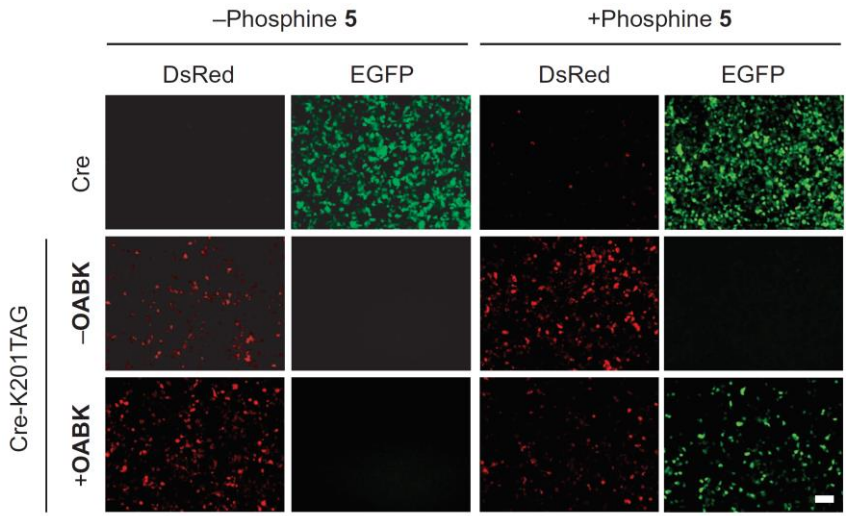
- Cre – A tyrosine recombinase from P1 bacteriophage recognizing *loxP* sites
- *loxP* site - ATAACTTCGTATA –[sequences(DsRed)]-TATACGAAGTTAT
- Design of the Mutation and the System for Demonstration

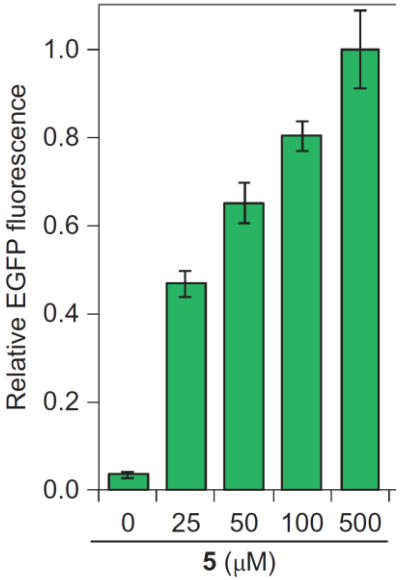


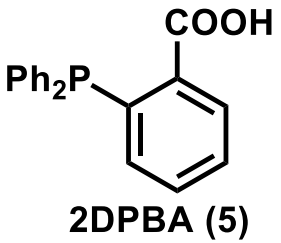


**K201 is essential for the interaction between Cre and DNA double strand*

- Results



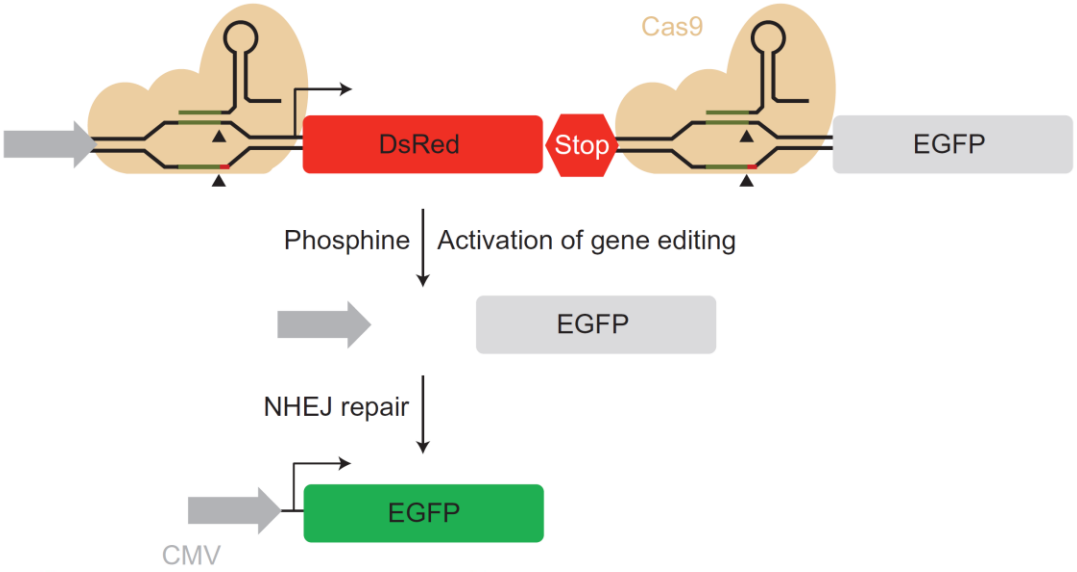




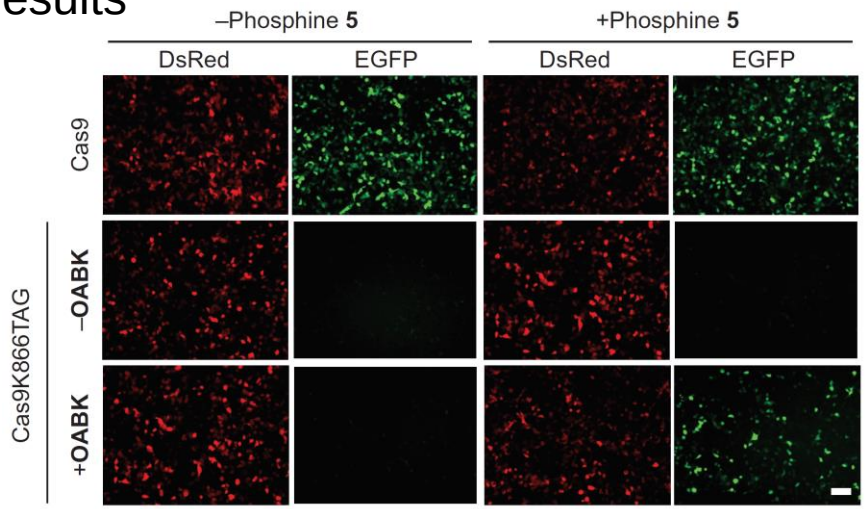
CRISPR/Cas9 Gene Editing

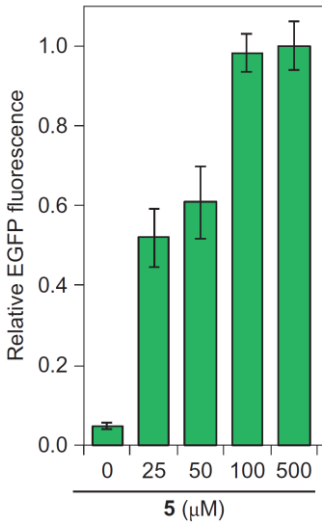
- Cas9 (CRISPR associated protein 9): an RNA-guided DNA endonuclease
- Design of the Mutation and the System for Demonstration

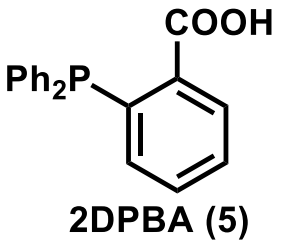
  *K866 of Cas9 is essential for interaction between Cas9 and guide-RNA



- Results







Summary

- Incorporation of Unnatural Amino Acids
  - Utilizing aaRSs from bacteria and archaea
  - Substrate scope has been being expanded
(up to present, Lys, Trp and Phe derivatives are major)
  - Usage is not difficult
- Protein Control with Bioorthogonal Ligand Tethering (BOLT)
  - Highly selective
  - Fast switching with light
  - On/off switching
  - A specific amino acid residue is not necessary for mutation
  - Difficult to design the mutation
  - Ligand might affect on protein function (it's not native)
  - Mother inhibitor is required
  - *In vivo* applicability?
- Protein Control with Staudinger Reduction
  - Highly selective
  - Versatile to control protein function
  - Easy to design the mutation
  - Producing the native protein
  - Lysine has to be involved in the function of target protein
  - Reaction is not fast
  - *In vivo* applicability?

Appendix-1: Coordinated Phe derivatives

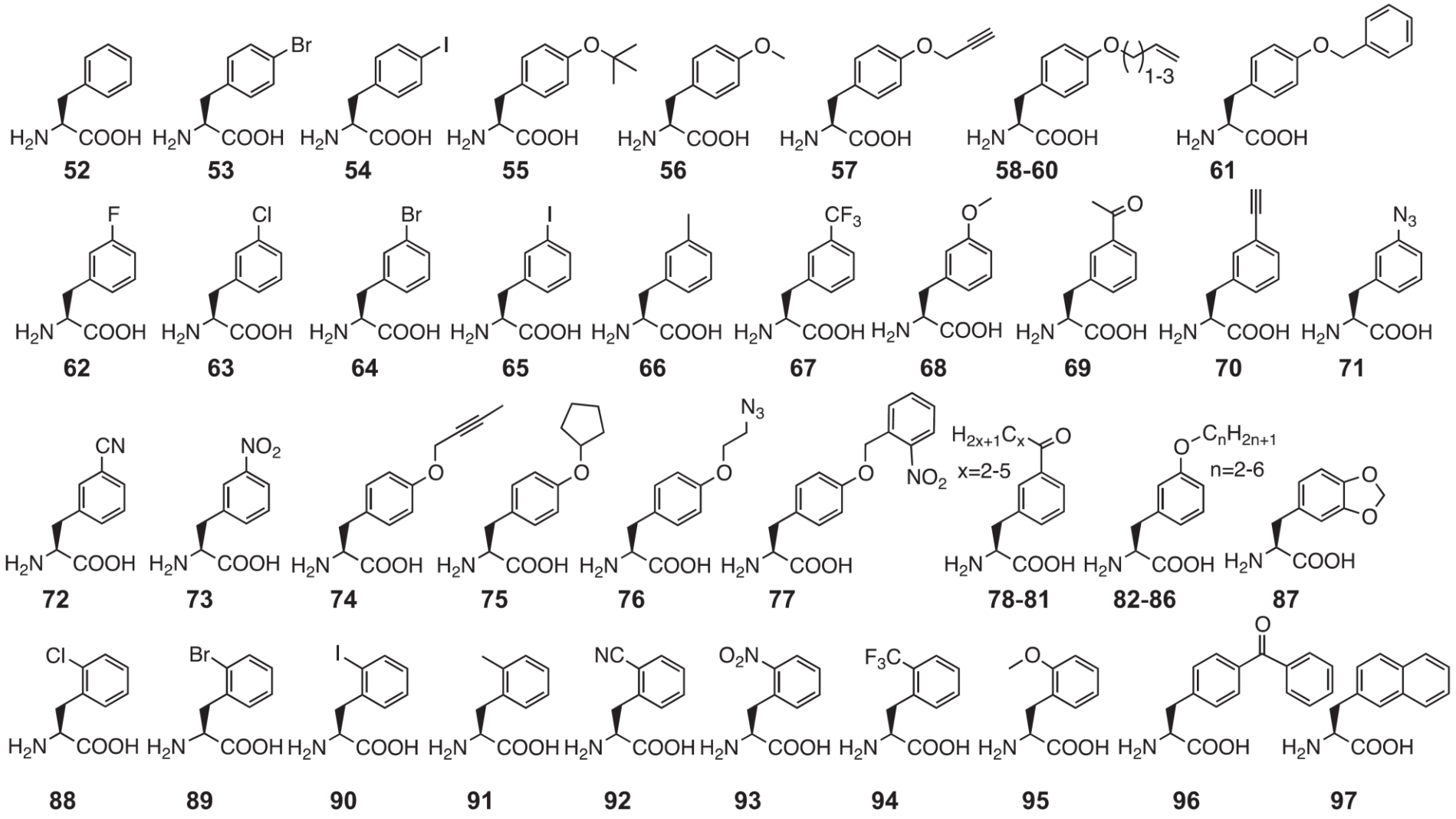
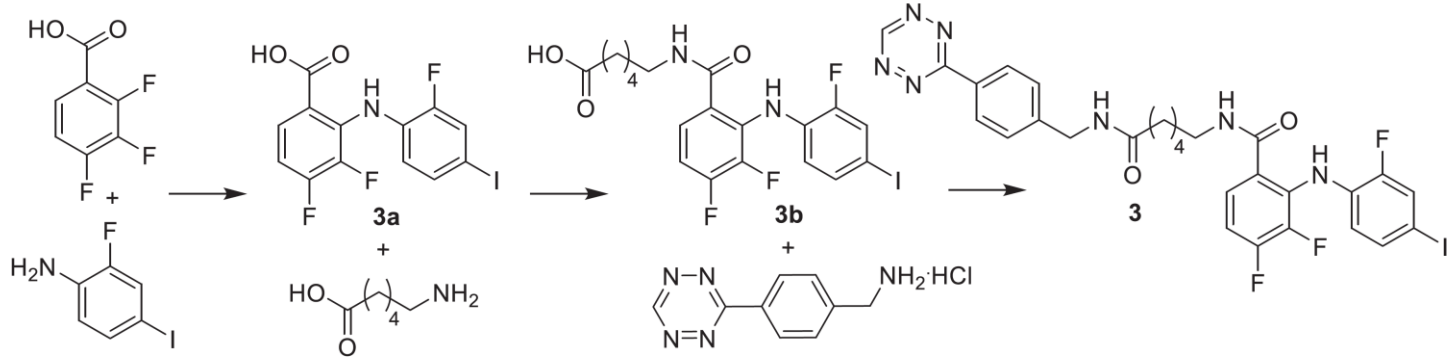


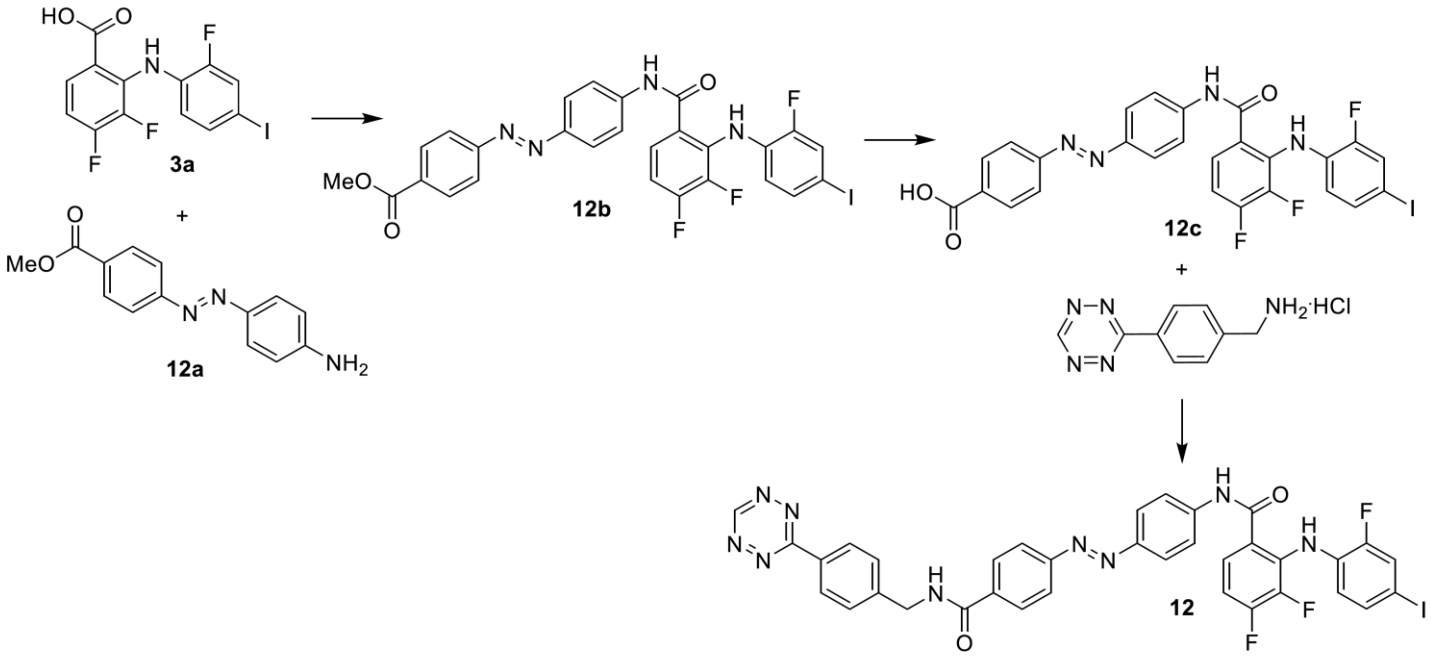
Fig. 9. Phe derivatives that have been genetically incorporated into proteins using engineering PylRS mutants in coordination with $tRNA^{Pyl}$.

Wan, W.; Tharp, J. M.; Liu, W. R. *Biochim Biophys Acta.* **2014**,*1844*, 1059

Appendix-2: Synthesis of iBOLT probe



Scheme S1. Synthesis of compound **3**.



Scheme S10. Synthesis of compound **12**.

Appendix-3: Synthesis of OABK

Synthesis of *ortho*-azidobenzyloxycarbonyl lysine (OABK) HCl salt