N-terminal specific protein modification

Literature seminar #1

2020.1.30

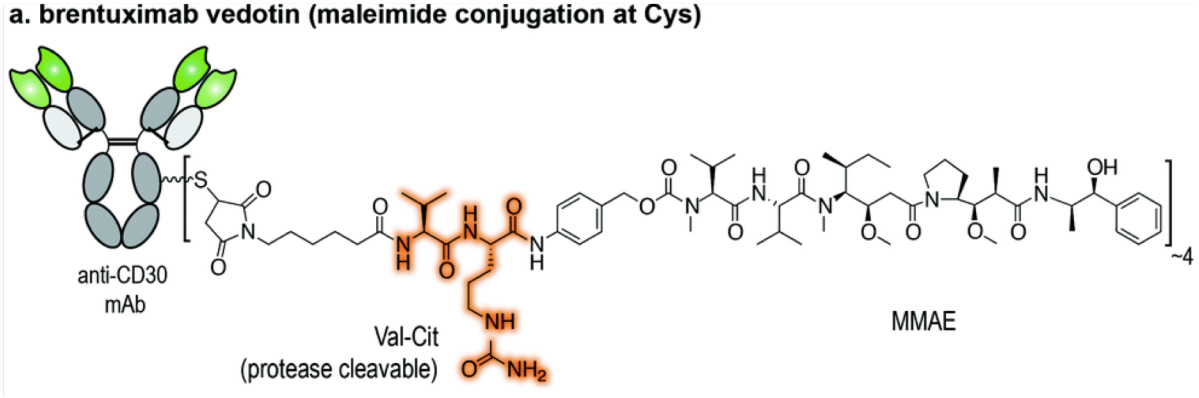
B4 Tamiko Nozaki

Contents

Contents

Protein Bioconjugation

✓ Cysteine side chain modification



・Inappropriate for proteins whose catalytic function or structural stability rely on the Cys.

✓ Lysine side chain modification

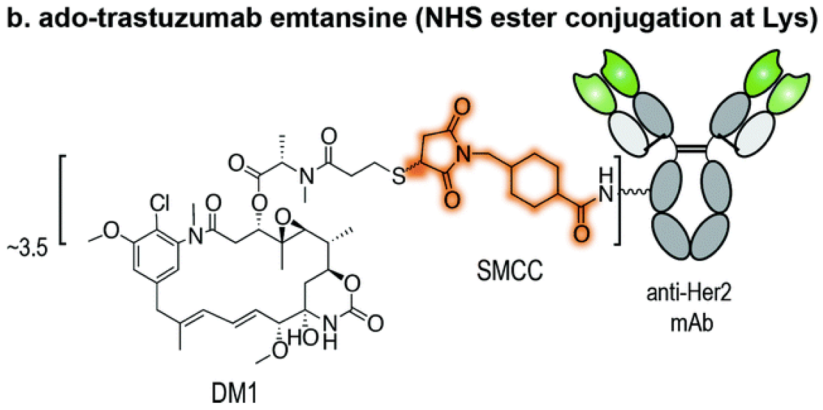


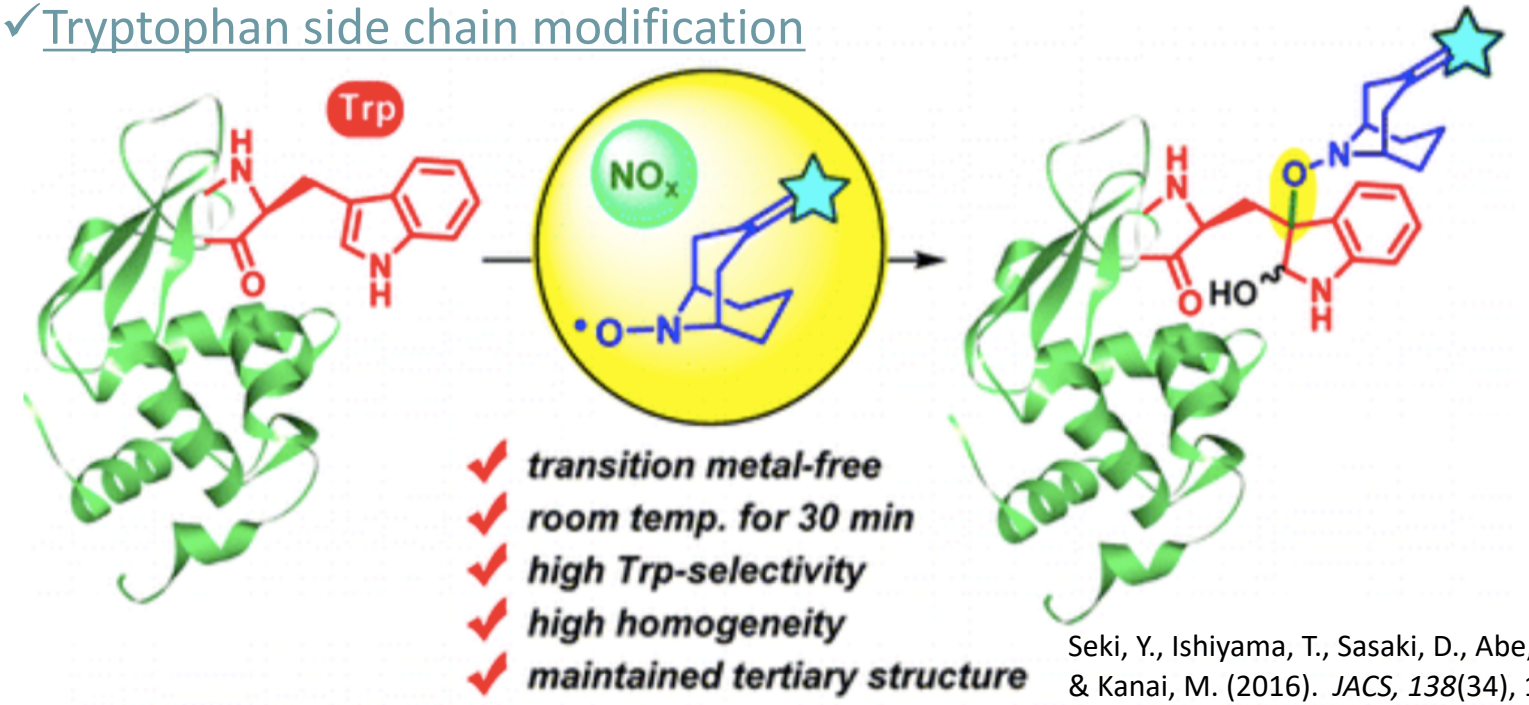
Fig. 1 Structures of currently FDA approved ADCs. (a) Brentuximab vedotin (Adcetris®; Seattle Genetics/Millennium Pharmaceuticals);[4] and (b) ado-trastuzumab emtansine (Kadcyla® – T-DM1; Roche/Genentech).[5] NHS ester: *N*-hydroxysuccinimide ester; Val-Cit linker: valine-citrulline linker; SMCC: succinimidyl-4-[*N*-maleimidomethyl]-cyclohexane-1-carboxylate; DM1: thiol-containing maytansinoid. MMAE: monomethyl auristatin E.

Akkapeddi, P., Azizi, S., Freedy, A. M., Cal, Pedro M. S. D, Gois, P. M. P., & Bernardes, G. J. L. (2016).. *Chemical Science, 7*(5), 2954-2963.

・Products are tend to heterogeneous mixture with altered activity.

Protein Bioconjugation

✓Tryptophan side chain modification



Seki, Y., Ishiyama, T., Sasaki, D., Abe, J., Sohma, Y., Oisaki, K., & Kanai, M. (2016). *JACS, 138*(34), 10798-10801.

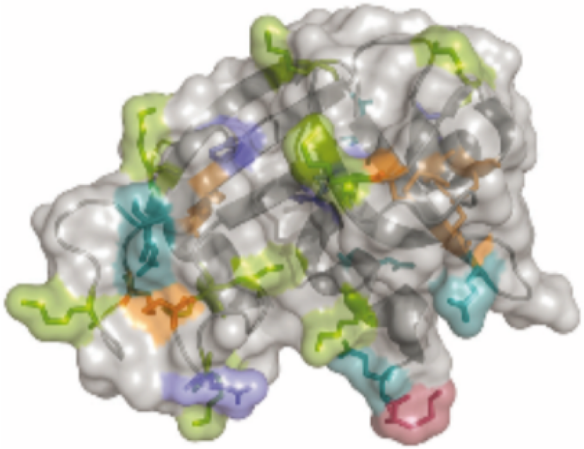
Trp : ・The least abundant (~1%)
・Little surface exposure
・Included in the primary sequences of 90% of native proteins

→Targeting Trp is the promising strategy to high selectivity while maintaining structural integrity.

✓ *N*-terminus is also the emerging target of of site-specific modification.

N-terminus as the target of protein modification

- α-amine group of the *N*-terminus is a uniquely reactive site for chemical modification in single-chain protein.
- *N*-terminal amine is charged as well as Lysine at physiological condition.
- 80.3 % of the *N*-terminal residues are solvent exposed .
- *N*-terminal modification has minimal effect for the overall protein structure.



	Avg. pK_a	Avg. abundance
Lysine	10.5	5.9%
Cysteine	8.0	1.9%
Aspartic acid	3.5	5.3%
Glutamic acid	4.0	6.3%
N terminus	6–8	Unique

Figure 1 | Schematic representation of the average abundance of the amino acids commonly used for bioconjugation, including the corresponding pK_a values of their side chains. The average (avg.) pK_a value of the N-terminal α-ammonium group is substantially lower than those of lysine ε-ammonium groups. This facilitates, but does not guarantee, site-specific modification at that position (RNAse A is shown; PDB ID: 2QCA).

Rosen CB, Francis MB. (2017), *Nature chemical biology*. 7;13(7):697-705.

✓ *N*-terminal residues are easily accessible, high selective modification target with little disruption to protein function.

pK_a of the α-amino group of N-terminal amino acid

Electrostatic interactions have significant effect on the pK_a of the α-amino group.

- α-carbonyl group
- *N*-terminal residue

Peptide analogue[a]	This study[b]	Free amino acid[c]
Pro (P)	7.1	10.6
Gly (G)	7.0	9.8
Asp (D)	6.8	9.6
Ala (A)	6.8	9.7
Glu (E)	6.6	9.7
Val (V)	6.5	9.6
Ile (I)	6.4	9.7
Gln (Q)	6.4	9.1
Trp (W)	6.3	9.4
Ser (S)	6.3	9.2
Thr (T)	6.3	9.1
Leu (L)	6.3	9.6
His (H)	6.3	9.2
Lys (K)	6.2	9.2
Asn (N)	6.1	8.8
Arg (R)	6.1	9.0
Tyr (Y)	6.1	9.1
Met (M)	6.1	9.2
Phe (F)	6.0	9.2
Cys (C)	–	10.5

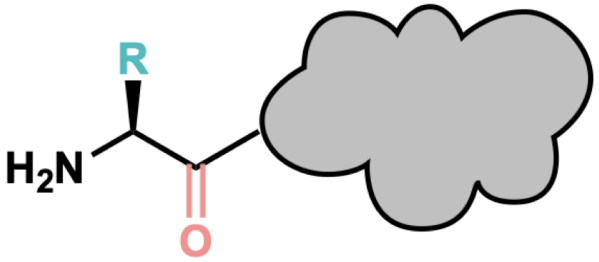


TABLE II

pK_a VALUES OF THE α-AMINO GROUPS OF PEPTIDE ANALOGUES CONTAINING 19 DIFFERENT N-TERMINAL AMINO ACID RESIDUES

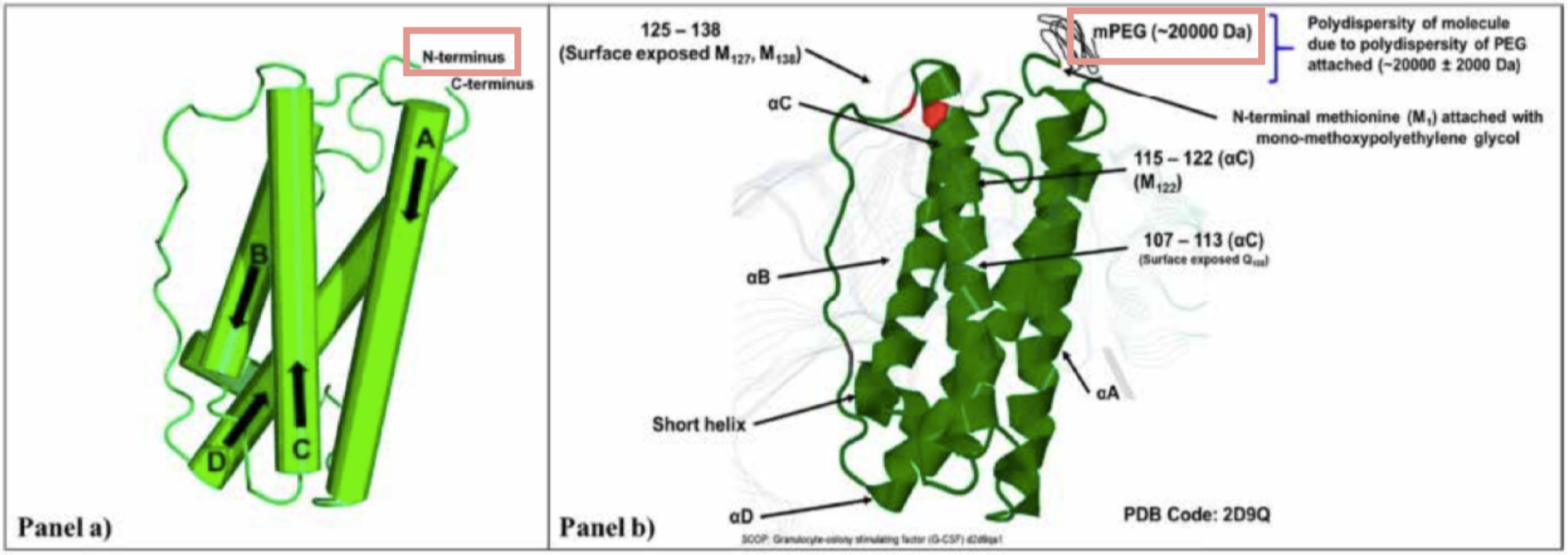
Sereda TJ, Mant CT, Quinn AM, Hodges RS.,(1993), Journal of Chromatography A,646(1):17-30.

✓ pK_a of α-amino group is around 6-8 (pK_a of ε-amino group of Lysine : 10.3)

N-terminal modification for therapeutics

Pegfilgrastim (Neulasta)

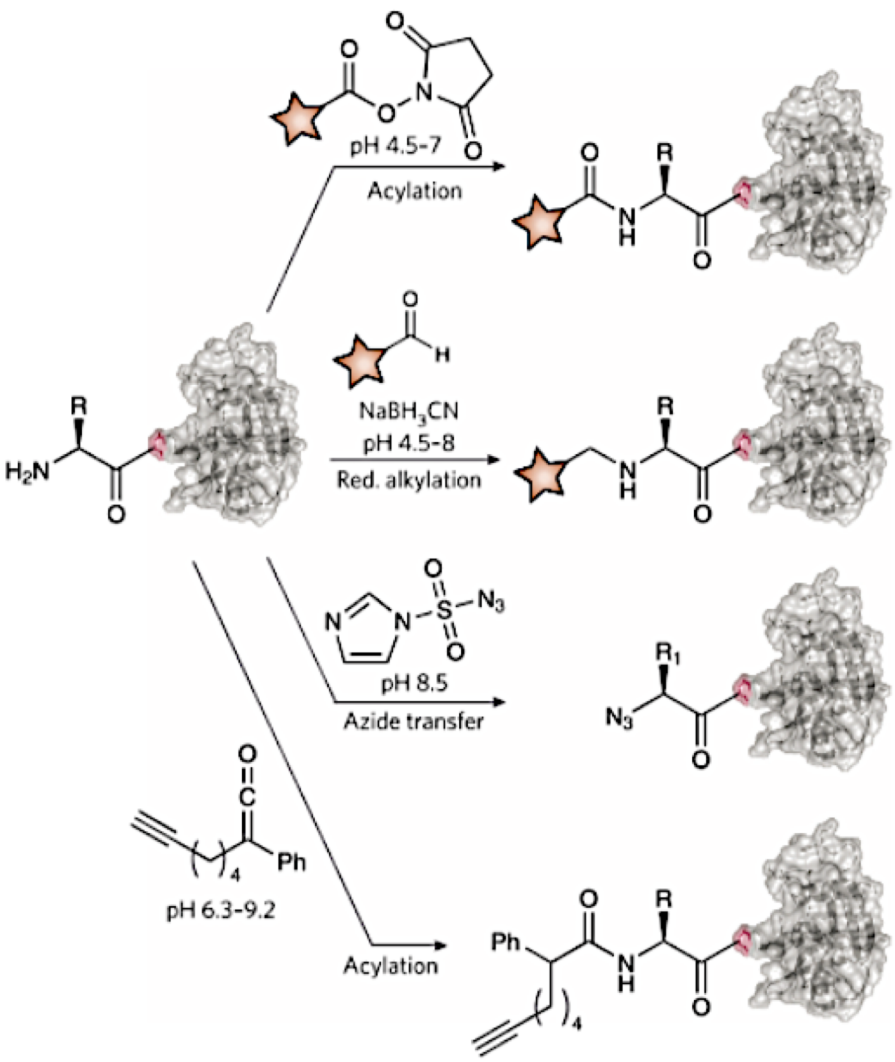
: N terminally PEGylated human recombinant granulocyte colony-simulating factor (rh-GCSF)



https://www.tga.gov.au/sites/default/files/auspar-pegfilgrastim-190328-pi.pdf

✓ Increasing the stability and circulation half time of protein without substantial effects on protein function and charge.

pH control method



(Rosen CB, Francis MB., 2017)

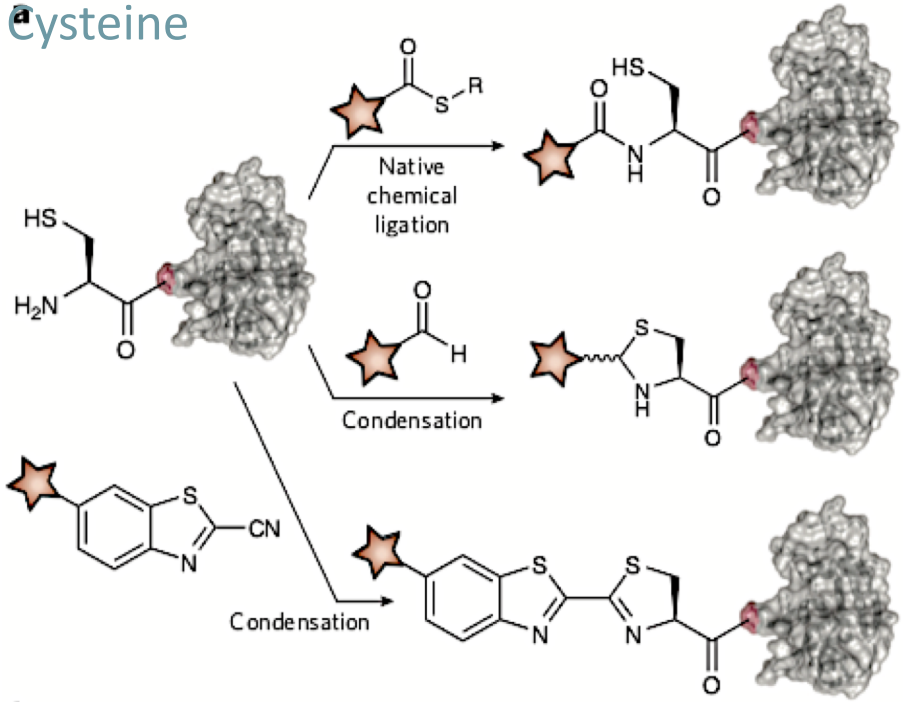
○ Applied to any *N*-terminal amino acids and the wide variety of attachment group.

● Complete site specificity is rarely achieved.

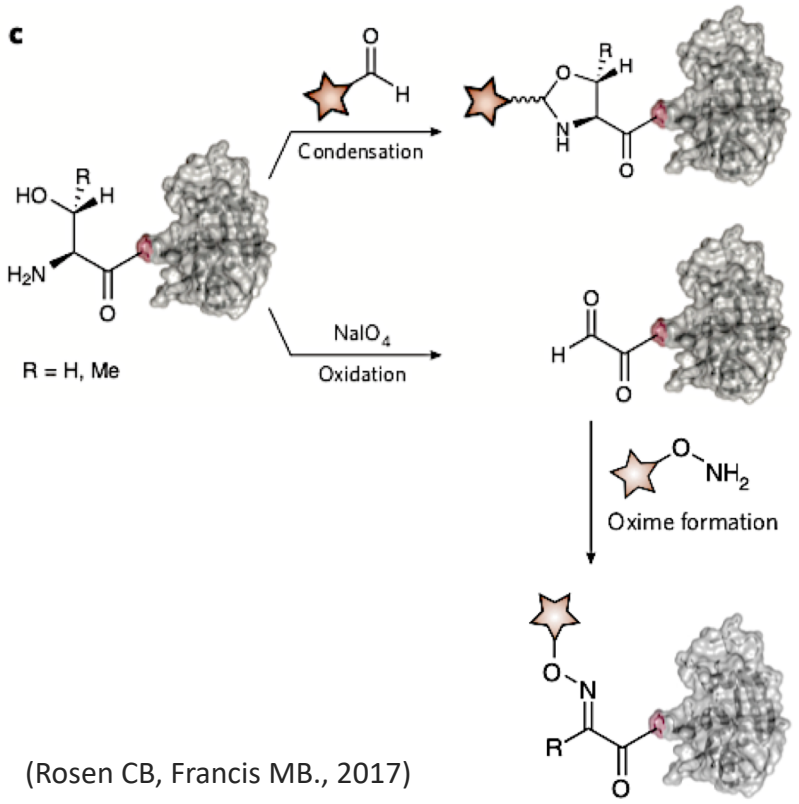
Figure 2 | Selective modification of protein N termini using pH control. Methods include (from the top) acylation using an activated ester, alkylation by reductive (red.) amination with aldehydes, azide incorporati with a diazotransfer reagent, and acylation using a ketene for the introduction of an alkyne.

Using Specific side chain amino acid

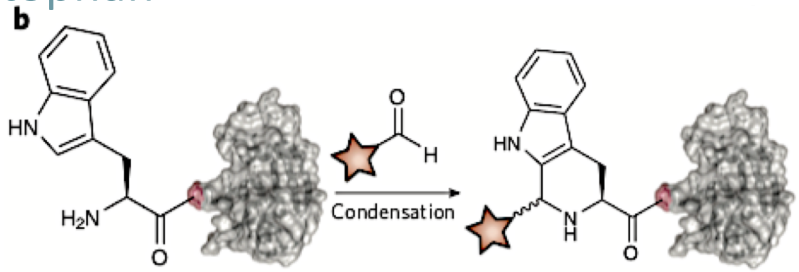
Cysteine



Serine



Tryptophan

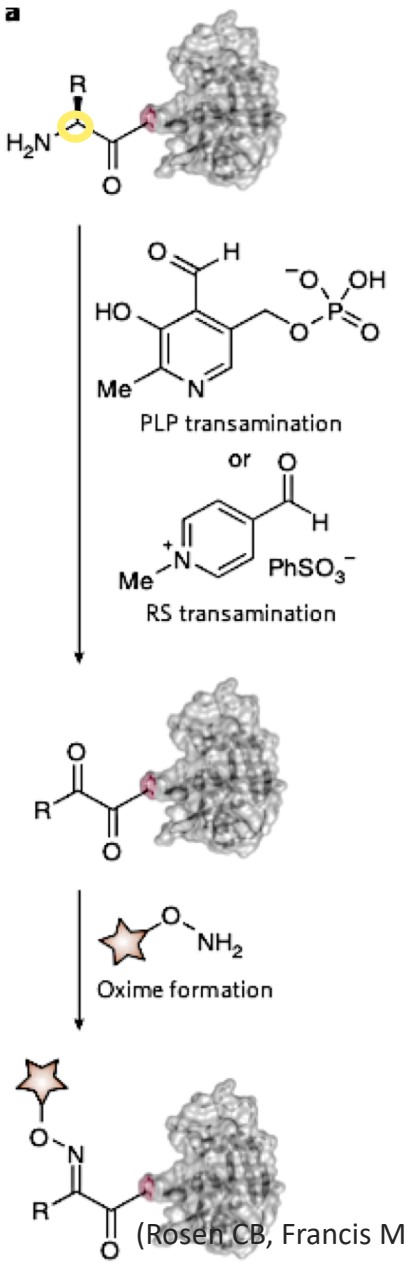


(Rosen CB, Francis MB., 2017)

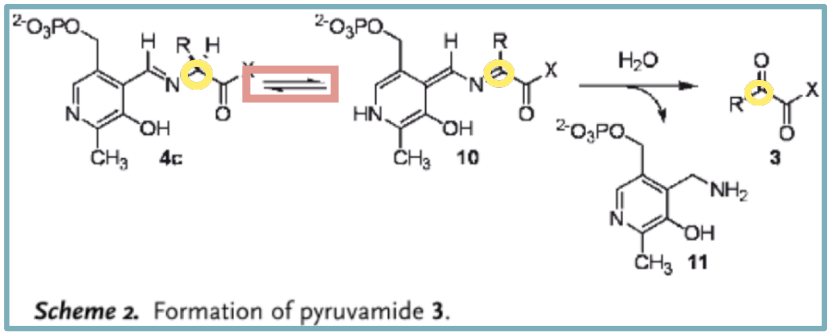
Figure 3 | Modification of specific amino acids at protein N termini. (**a**) N-terminal cysteine modification through native chemical ligation with thioesters, condensation with aldehydes and condensation with 2-cyanobenzothiazole derivatives. (**b**) Pictet–Spengler reactions, in which tryptophan residues form cyclic products with aldehydes. (**c**) Functionalization of N-terminal serine or threonine residues by condensation with aldehydes to form oxazolidines or through oxidation with sodium periodate ($NaIO_4$) to generate an aldehyde that can be further modified by oxime formation with alkoxyamines.

- These methods need introduction of specific amino acid at *N*-terminus.

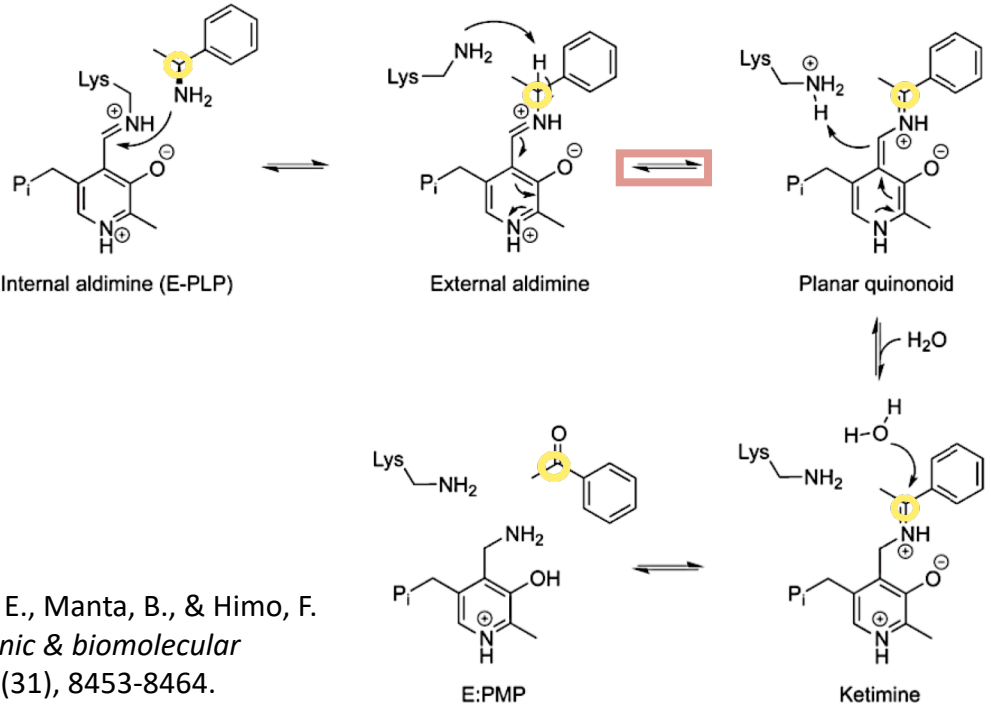
N-terminal transamination



(Rosen CB, Francis MB., 2017)



Gilmore, J. M., Scheck, R. A., Esser-Kahn, A. P., Joshi, N. S., & Francis, M. B. (2006). *ACIE, 45*(32), 5307-5311.



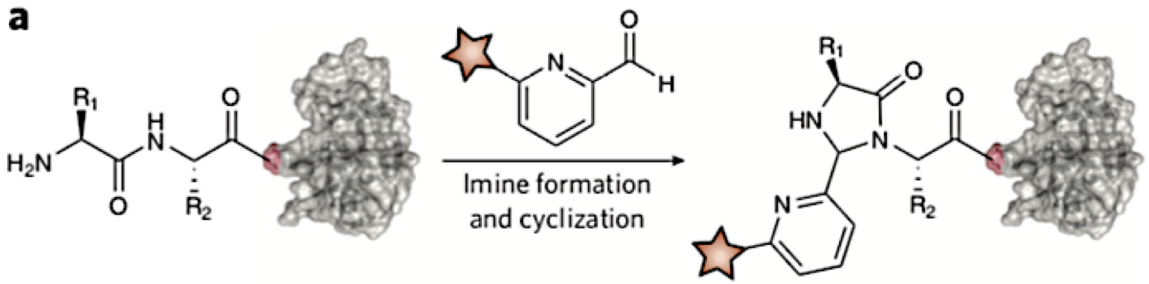
Cassimjee, K. E., Manta, B., & Himo, F. (2015). *Organic & biomolecular chemistry*, *13*(31), 8453-8464.

Scheme 1 Generally suggested half-transamination reaction where the (*S*)-1-phenylethylamine amino donor is converted to the corresponding ketone, acetophenone, while PLP is converted to PMP.

- AKT sequence at the *N*-terminus is optimal.

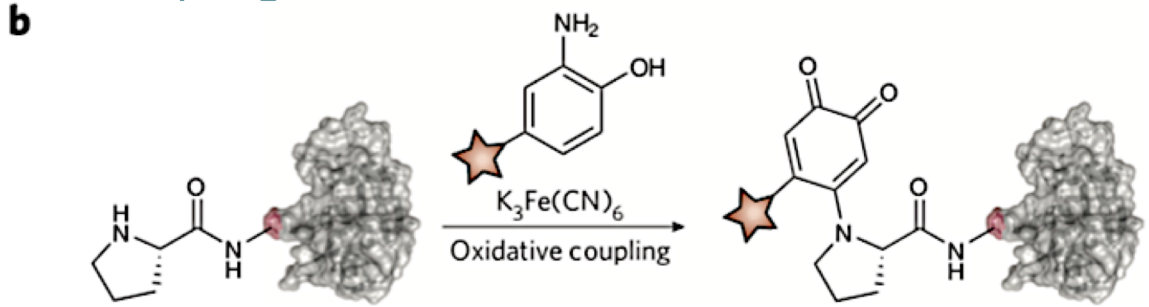
Other one step chemical modification

Imidazolidinone formation



- Potentially general approach.
- The absence of a proline in the second position is necessary.

Oxidative coupling

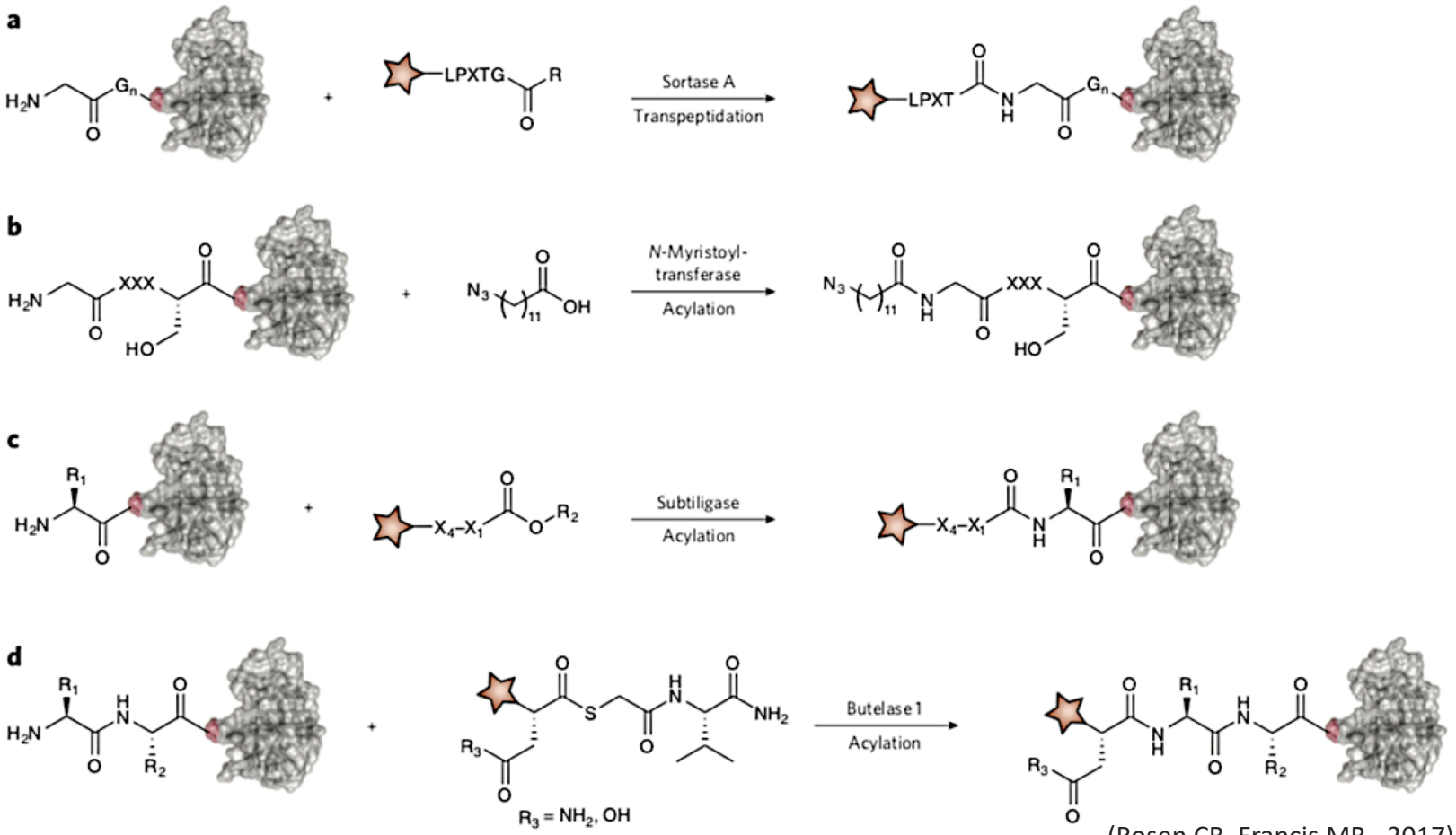


(Rosen CB, Francis MB., 2017)

Figure 5 | One-step modification of protein N termini. (**a**) Imidazolidinone formation with 2-pyridinecarboxaldehyde (2PCA) derivatives. (**b**) Oxidative coupling of *o*-aminophenols to N-terminal proline residues in the presence of potassium ferricyanide ($K_3Fe(CN)_6$).

- Free Cys are reacted with *orhto*-aminophenol.

N-terminal modification by enzymes



(Rosen CB, Francis MB., 2017)

Figure 7 | Enzyme-mediated N-terminal protein modification. (**a**) Sortase A (SrtA)-catalyzed transpeptidation of an LPXTG peptide derivative for attachment to an H_2N-$(G)_n$-protein (X can be any amino acid). (**b**) Acylation of an H_2N-GXXXS/T-protein with an azide-containing fatty acid using *N*-myristoyltransferase (NMT) (X can vary). (**c**) Attachment of a glycolate ester substrate mediated by subtiligase. R_1, R_2, and X_4–X_1 can vary, and the residues X_4 and X_1 dominate substrate specificity. (**d**) N-terminal acetylation with an Asn/Asp thiodepsipeptide using butelase 1. Here R_1 can be any amino acid side chain except Pro, and R_2 is Ile, Val, Leu or Cys.

◯ The reactions occur under mild condition.

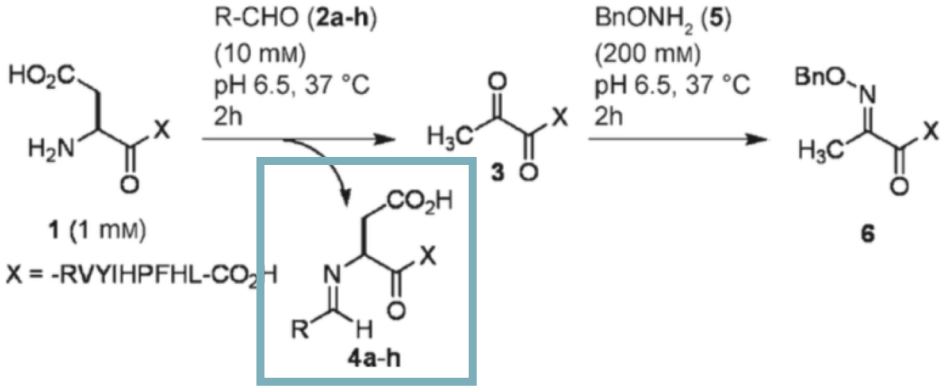
Summary of *N*-terminus selective protein modification methods

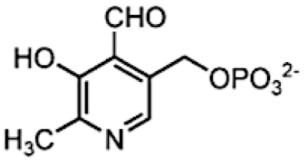
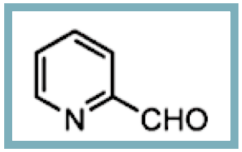
	Independence of amino acid sequence	The accessibility of the reagent	Site selectivity	High yield
pH control	○	◎/○	△	○
Using Specific side chain amino acid	×	○/△	○/△	-
transamination	○	△	○	-
Imidazolidinone formation	○	○	○	-
Oxidative coupling	×	△	△	◎
enzymes	△	△	○	-

→ Imidazoline formation method would be the general and selective method for N-terminal modification.

Contents

Identification of 2PCA



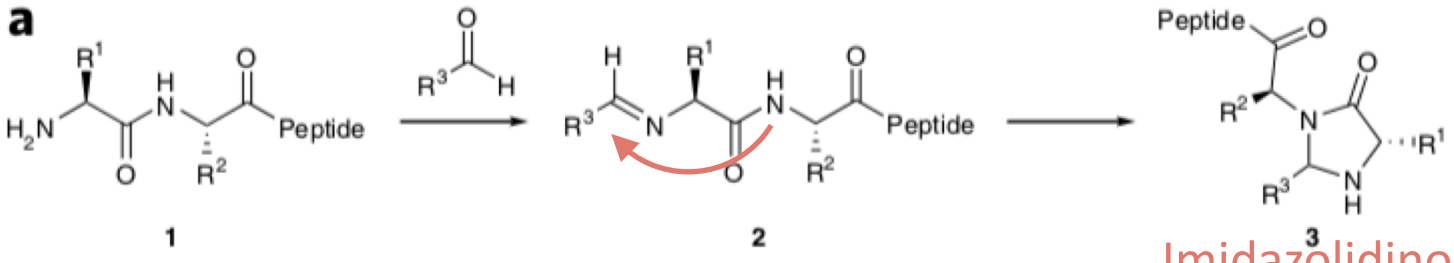
Aldehyde	3 [%]	4a–h [%]
	65	0
2g 	trace	> 90

Scheme 1. Formation of oxime 6 from the N-terminal aspartic acid residue of angiotensin I (**1**). Bn = benzyl.

[a] Conditions: 1 mM **1**, 10 mM aldehyde, 50 mM phosphate buffer (pH 6.5), 37 °C, 2 h. Product distributions were determined by ESI-MS.

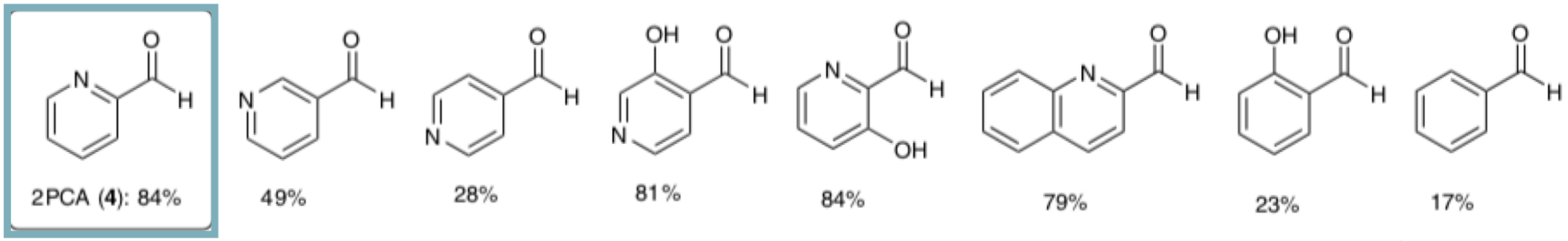
Gilmore, J. M., Scheck, R. A., Esser-Kahn, A. P., Joshi, N. S., & Francis, M. B. (2006). *ACIE, 45*(32), 5307-5311.

Imidazolidinone formation



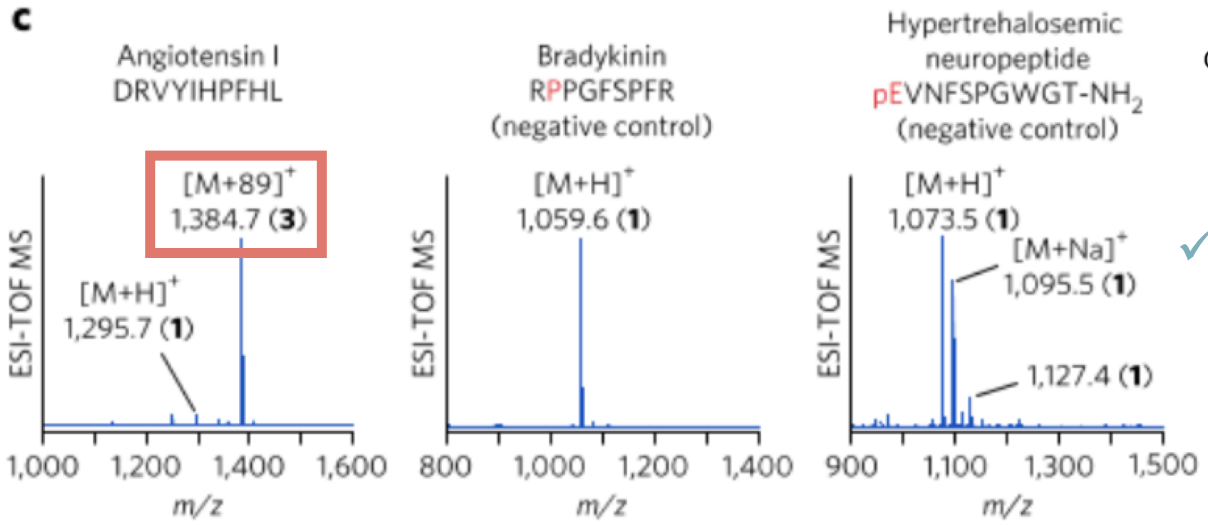
Optimization of aldehyde

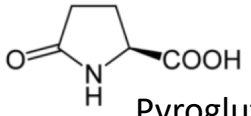
MacDonald JI, Munch HK, Moore T, Francis MB.(2015). *Nature chemical biology*;11(5):326-331.



Verification of *N*-terminus selective modification by 2PCA

✓ Imidazoline formation is *N*-terminal selective.



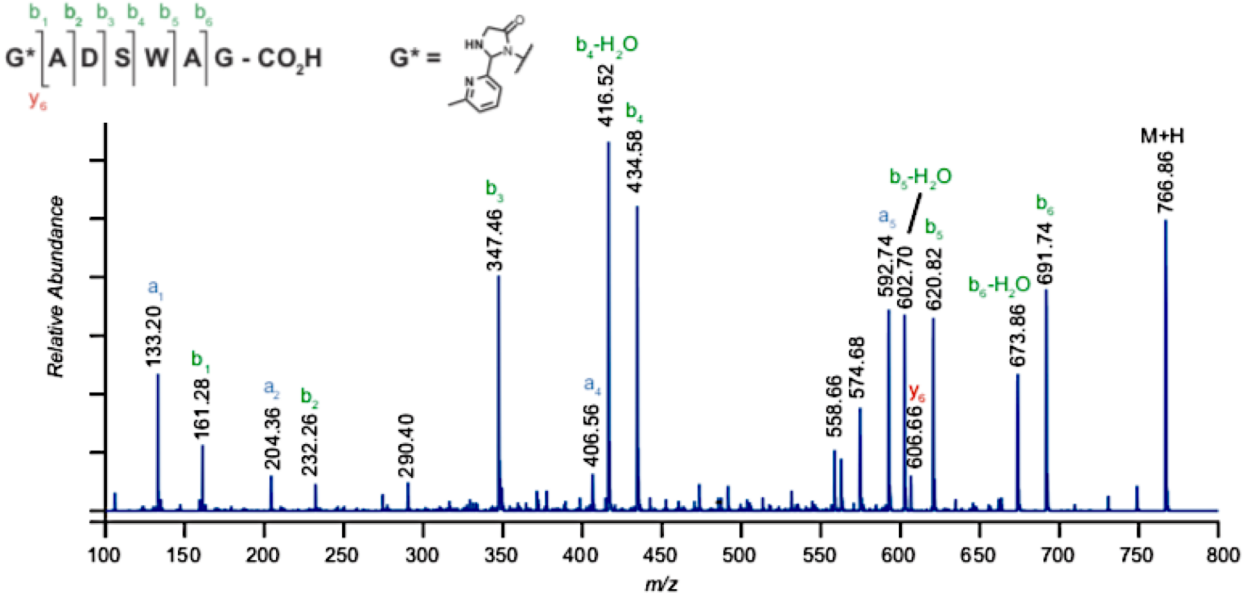


Pyroglutamic acid

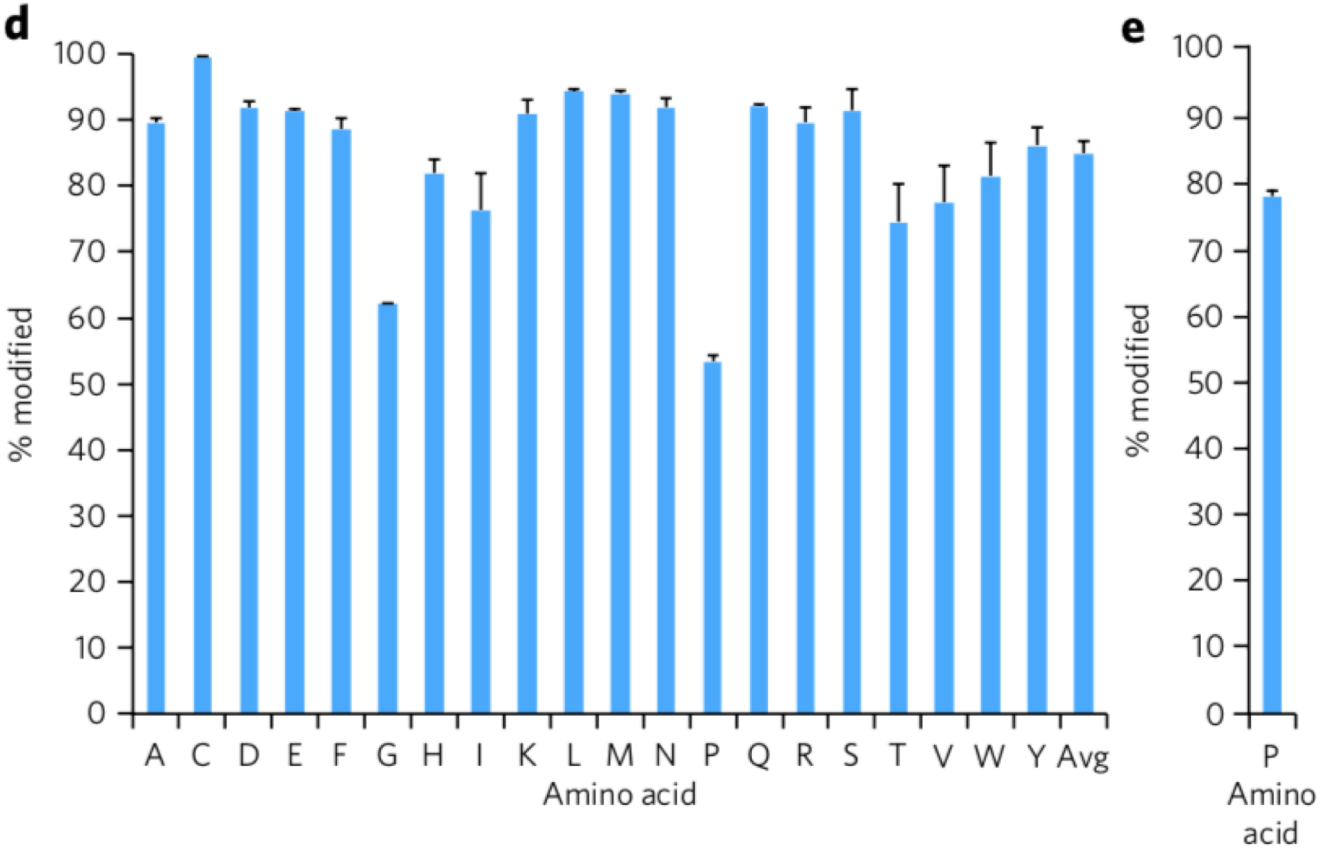
✓ Modification does not occur when N terminal is blocked or the second amino acid is Pro.

(MacDonald JI *et al*, 2015)

LC-MS/MS analysis



Reactivity of various *N*-terminal amino acids with 2PCA



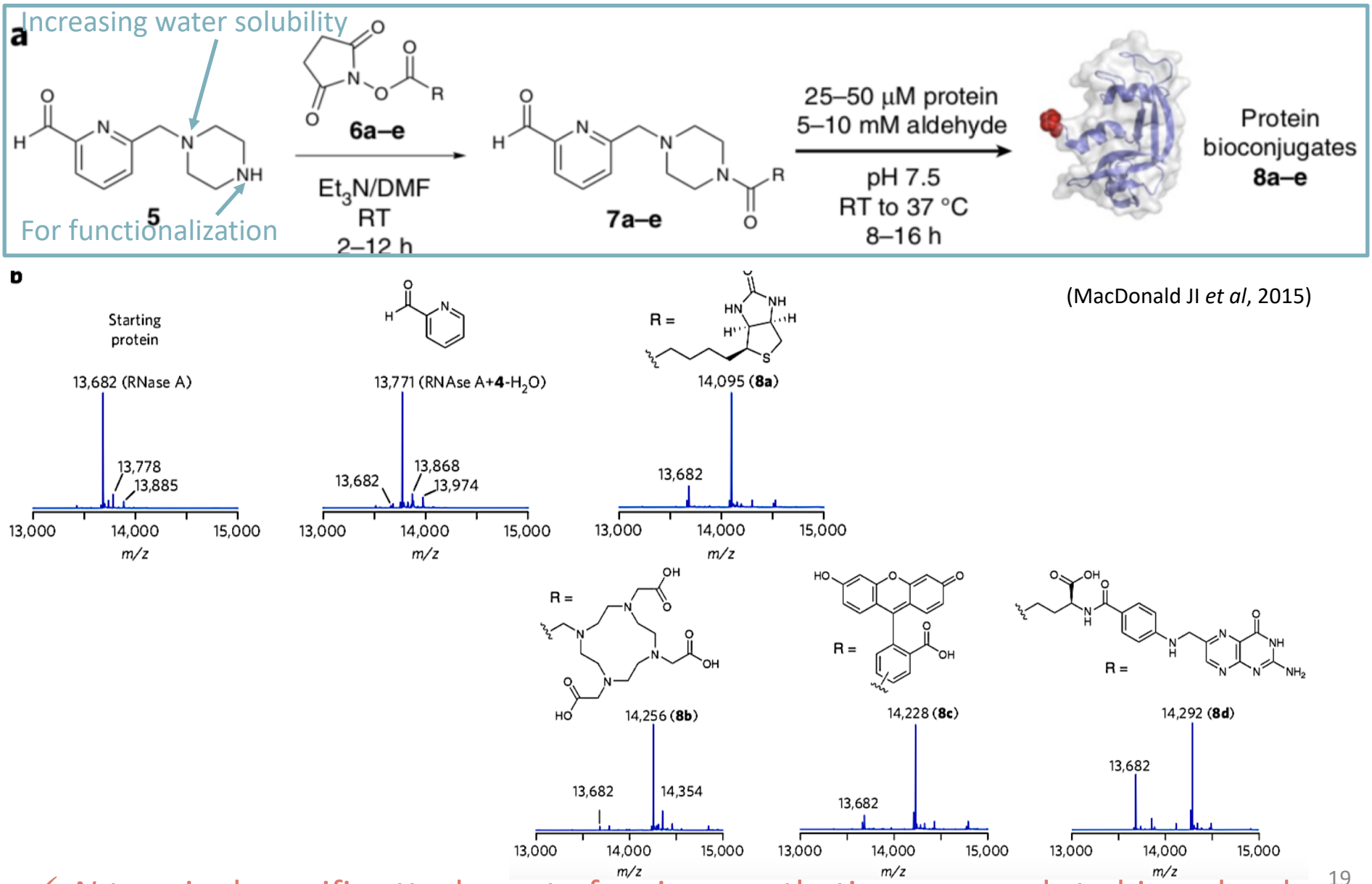
(MacDonald JI *et al*, 2015)

(**d**) The reaction showed good-to-excellent levels of conversion as quantified by LC/MS for a panel of X-ADSWAG peptides, where X represents a variable amino acid position. Conditions for **b**–**d**: 100 μM peptide, 10 mM 2PCA (or derivative), 10 mM phosphate buffer at pH 7.5 and 37 ° C. The samples were incubated for 1 h in **b** and for 4 h in **c** and **d**.

(**e**) The PADSWAG peptide was subjected to similar conditions as those in **d** but at pH 8.5. For **d** and **e**, experiments were run in triplicate, and error bars represent the s.d.

✓ All of the *N*-terminal amino acid residues provided relatively good yields.

Bioconjugation of diverse functional group

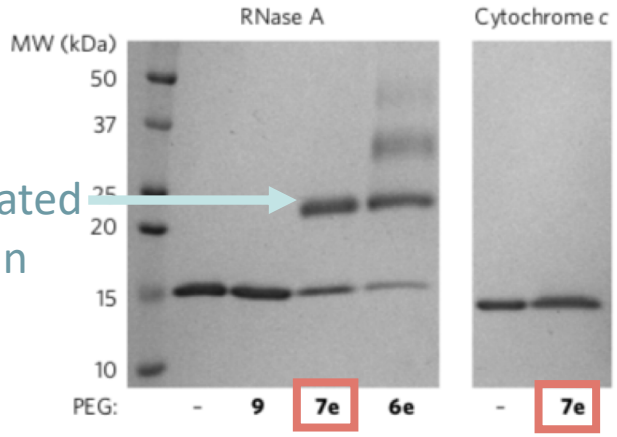


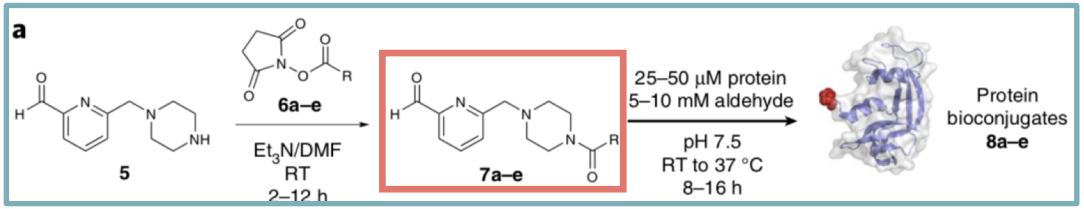
(MacDonald JI *et al*, 2015)

✓ *N*-terminal specific attachment of various synthetic compounds to biomolecules.

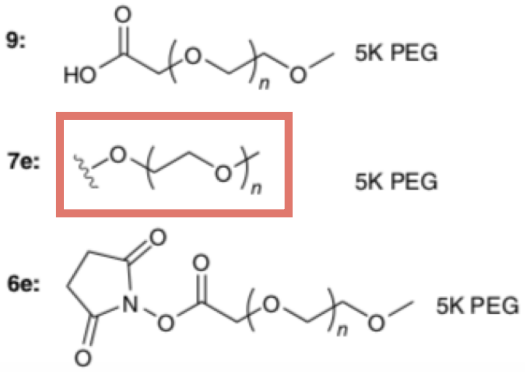
N-terminal selective bioconjugation with 2PCA

Single PEGylated protein





(MacDonald JI *et al*, 2015)



→No PEGylation (negative control)

→Single PEGylation

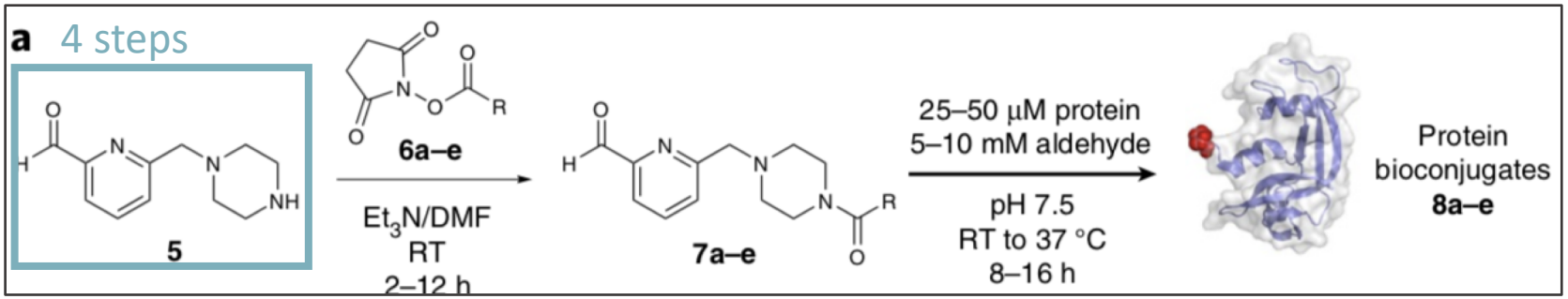
→heterogeneous PEGylation

2PCA :

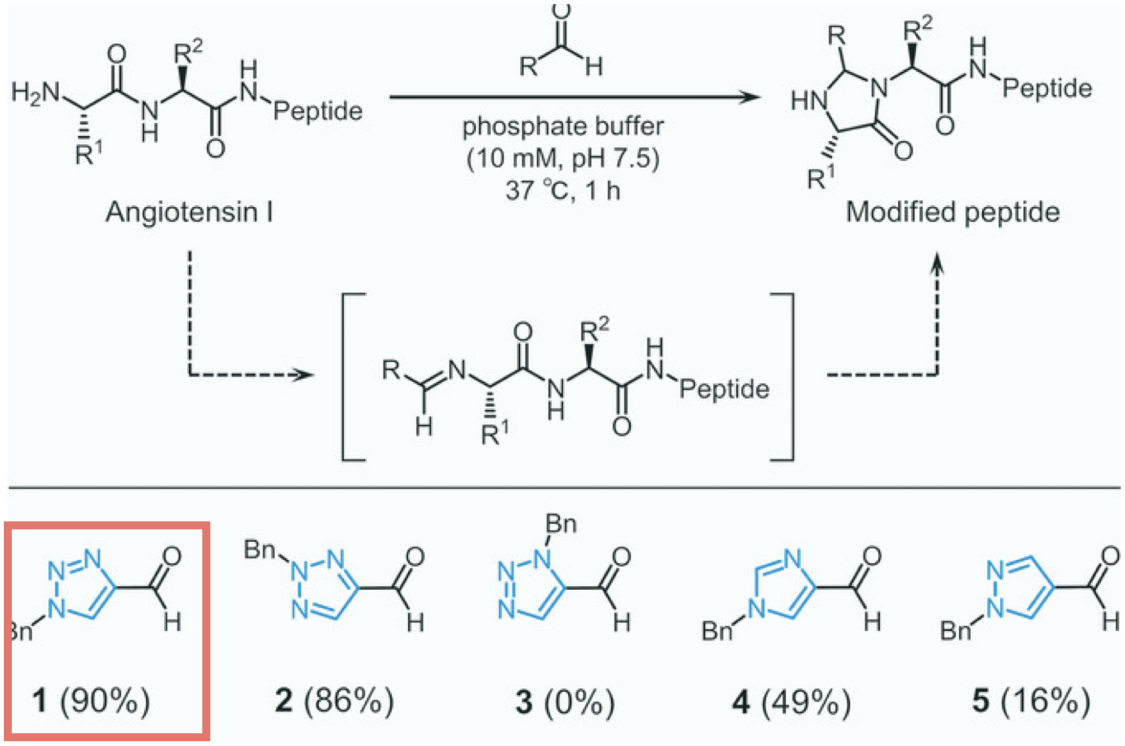
One-step selective *N*-terminal modification method under mild condition without genetic engineering

Identification of TA4C

2PCA



TA4C



Onoda, A., Inoue, N., Sumiyoshi, E., & Hayashi, T. (2020).. *ChemBioChem,*

✓ TA4C would be the one-pot synthesized *N*-terminal selective modification reagent.

Verification of *N*-terminus selective modification by TA4C

✓ *N*-terminal selective Imidazoline formation was confirmed.

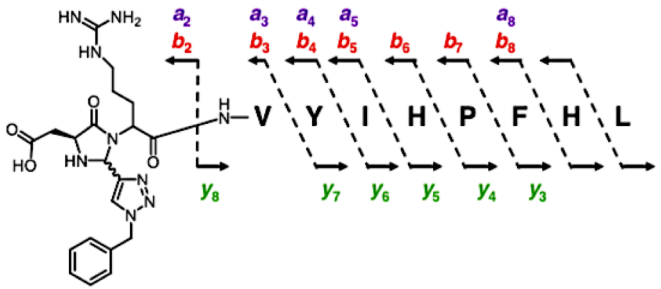
Table 1. Modification of peptides with reagent **1**.[a]

Entry	Peptide	Modification [%][b]
1	DRVYIHPFHL (angiotensin I)	92
2	WAGGSASGE (delta sleep-inducing peptide)	91
3	YGGFMRRVGRPE (BAM-12P)	96
4	LRQFLQKSLAAAA-NH_2 (neuronostatin-13)	98
5	**Ac**-YVG	–
6	**pE**AKSQGGSN (serum thymic factor)	–
7	R**P**PGFSPFR (bradykinin)	–

[a] Conditions: peptide (100 μM), **1** (10 mM), phosphate buffer (10 mM, pH 7.5) at 37 °C for 4 h. [b] The relative ratio of modification was analyzed by LC–MS measurements and used to determine the conversion.

✓ Modification does not occur when *N*-terminal is blocked or the second amino acid is Pro.

(Onoda *et al*, 2019)



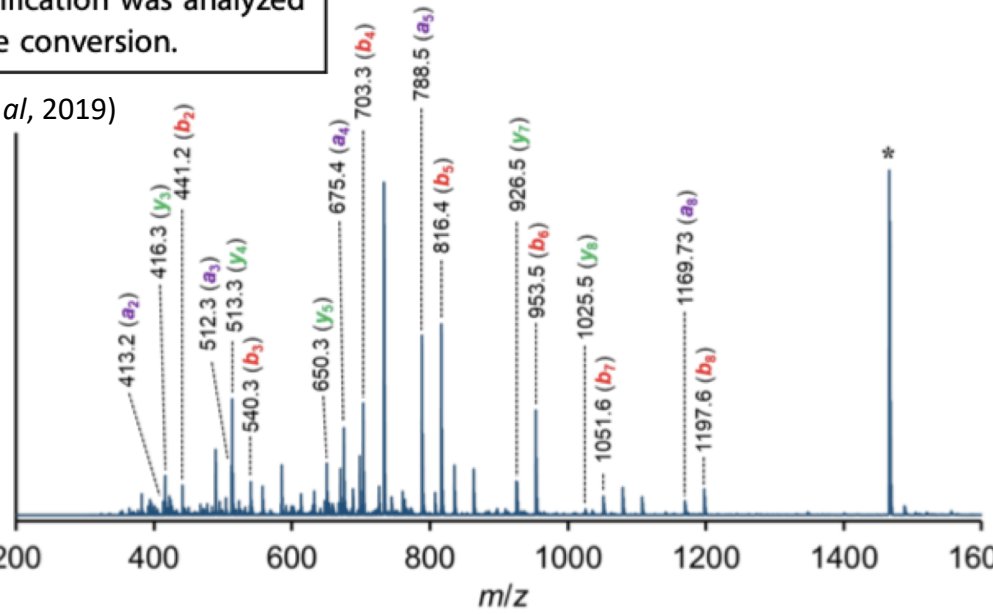
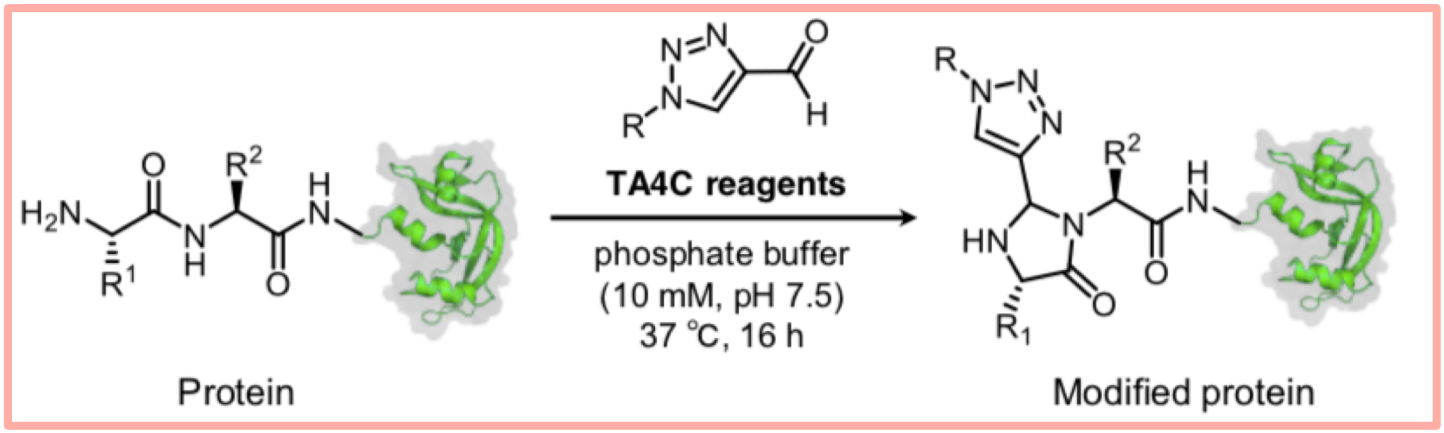


Figure S4. MS/MS spectrum of angiotensin I modified with 1. Observed a-, b- and y-ions are shown. The parent ion peak is labeled with an asterisk. Conditions: peptide (100 □M), 1 (10 mM), phosphate buffer (10 mM, pH 7.5) at 37 ° C for 4 h.

Bioconjugation of diverse functional group

Optimized condition



(Onoda *et al*, 2019)

b) Scope of functional molecules

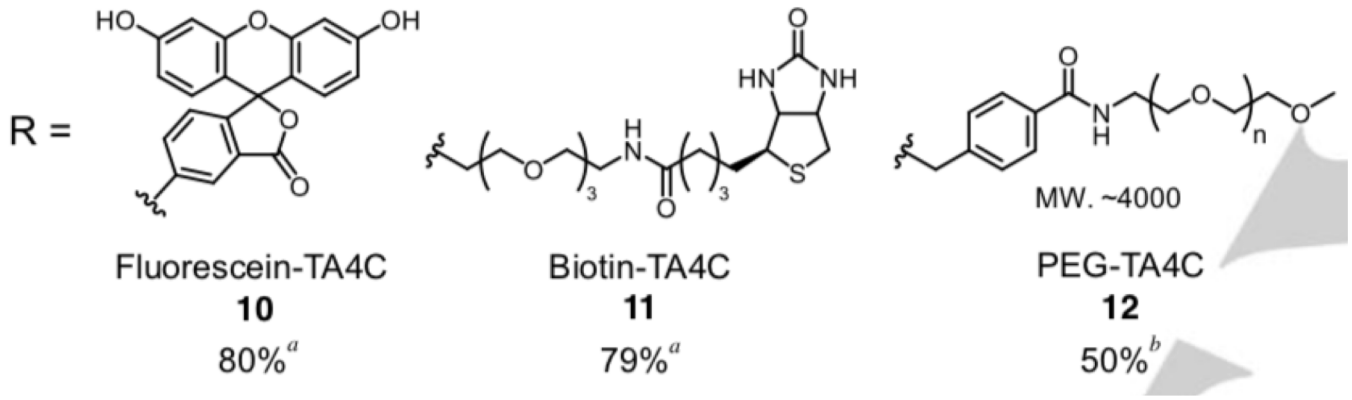


Figure 3. (b) Chemical structures of functional molecules attached to a TA4C moiety and relative ratio of modification of RNase. Conditions: protein (50 μM), 1, 10, 11 (10 mM) or 12 (20 mM), phosphate buffer (10 mM, pH 7.5) at 37 ° C for 16 h.

✓ *N*-terminal specific attachment of various functional molecules to protein

Facile process for generation of TA4C reagents

Dimroth rearrangement : 1-(4-nitrophenyl)-1H-1,2,,3-triazole-4-carbaldehyd

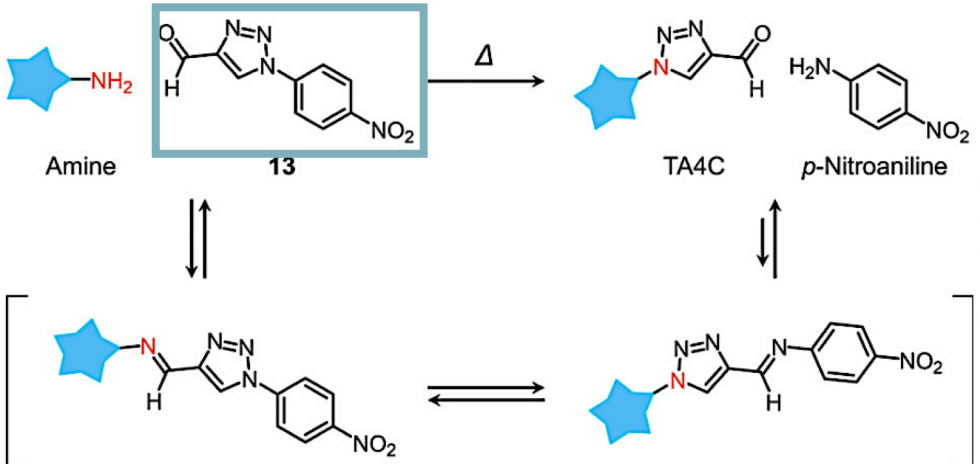
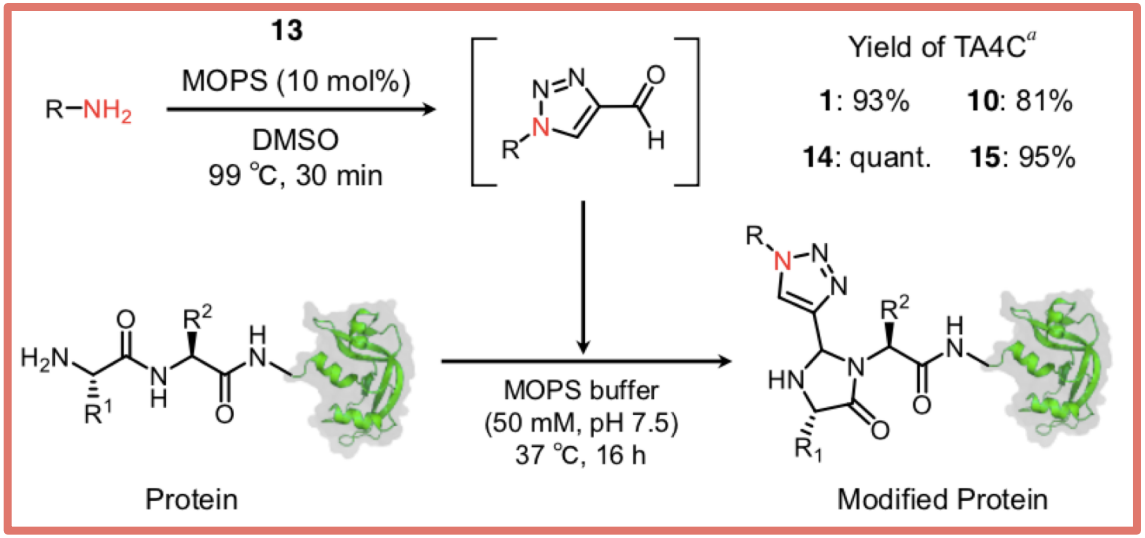


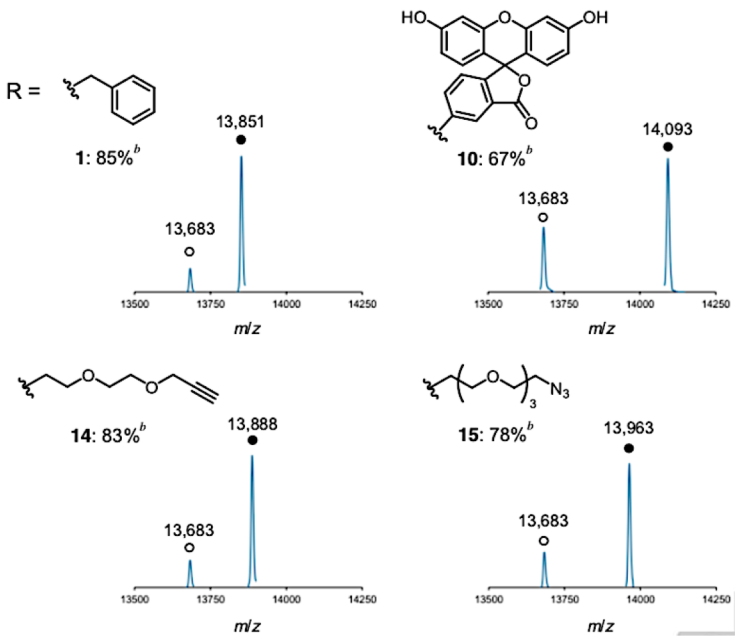
Figure 4. Dimroth rearrangement replacing 1-(4-nitrophenyl)-1*H*-1,2,3-triazole-4-carbaldehyde (**13**) with a labeling molecule having an amino group.

Figure 5. N-terminal modification of RNase using TA4C reagents prepared in a one-step process via a Dimroth rearrangement. Deconvoluted ESI-MS spectra of modified proteins are shown. Peaks for unmodified RNase are marked with open circles and species corresponding to the correctly modified protein are labeled with filled circles. Conditions for Dimroth rearrangement: amine precursor (100 mM), **13** (100 mM), MOPS (10 mol%), DMSO at 99 °C for 30 min. Conditions for protein modification: protein (50 μM), **1**, **10**, **14**, **15** (about 10 mM), MOPS buffer (50 mM, pH 7.5) at 37 °C for 16 h. [a] The conversion of Dimroth rearrangement was analyzed by ^{1}H NMR measurements. [b] The relative ratio of modification was analyzed by LC–MS measurements.

(Onoda *et al*, 2019)

Optimized condition





✓ Two step *N*-terminal modification

TA4C reagent with azide/alkyne group for chemical modification

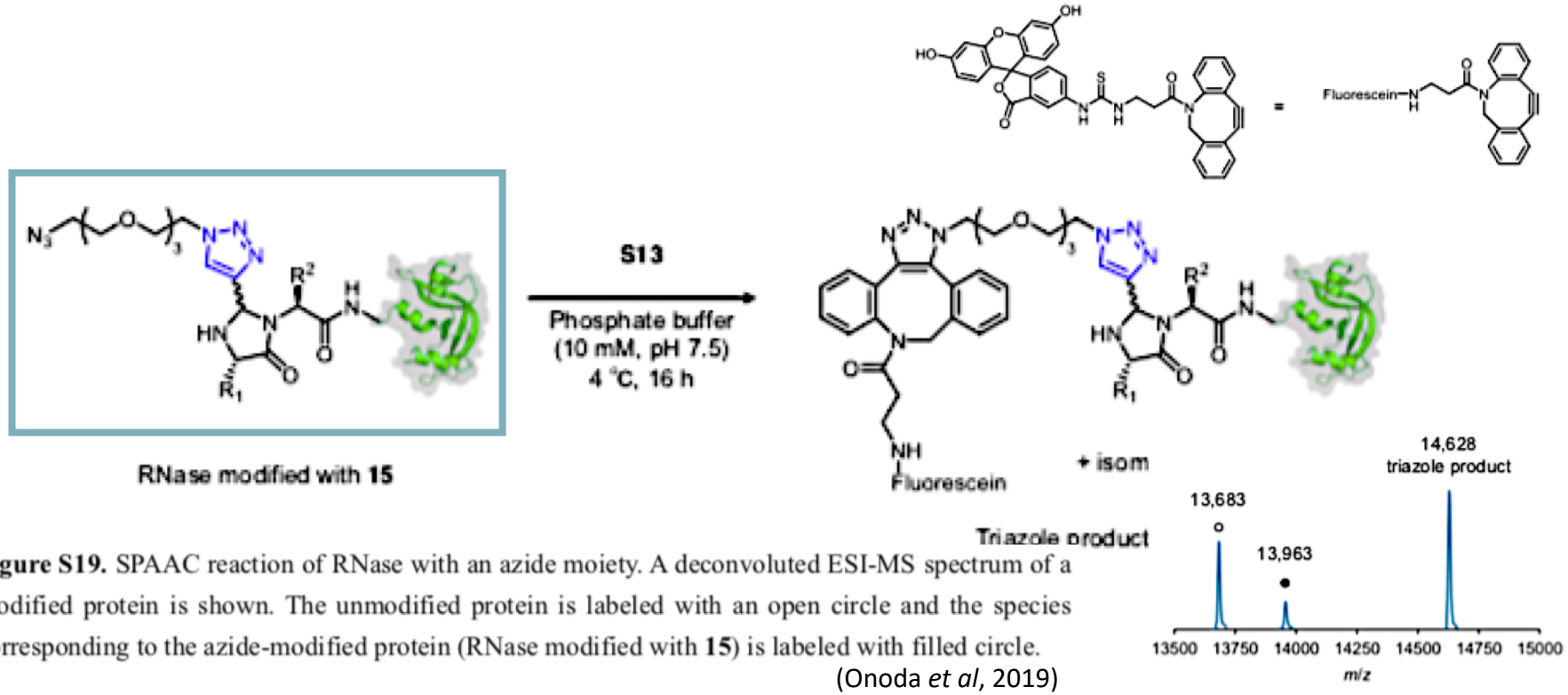


Figure S19. SPAAC reaction of RNase with an azide moiety. A deconvoluted ESI-MS spectrum of a modified protein is shown. The unmodified protein is labeled with an open circle and the species corresponding to the azide-modified protein (RNase modified with **15**) is labeled with filled circle.

(Onoda *et al*, 2019)

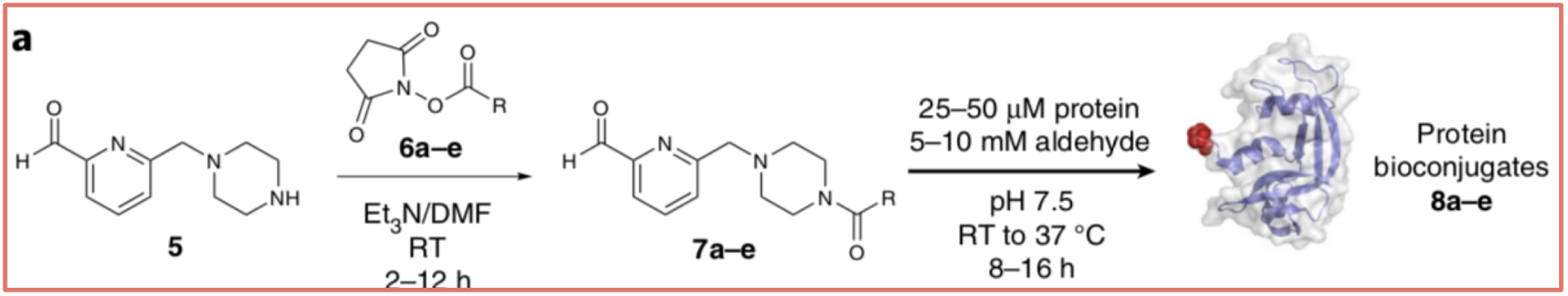
Dimroth rearrangement enables to prepare TA4C reagent with azide or alkyne group.
→ Strain-promoted azide-alkyne cycloaddition with functional moiety.

TA4C :

One step *N*-terminal protein modification reagent synthesized in CuAAC reaction or Dimroth rearrangement

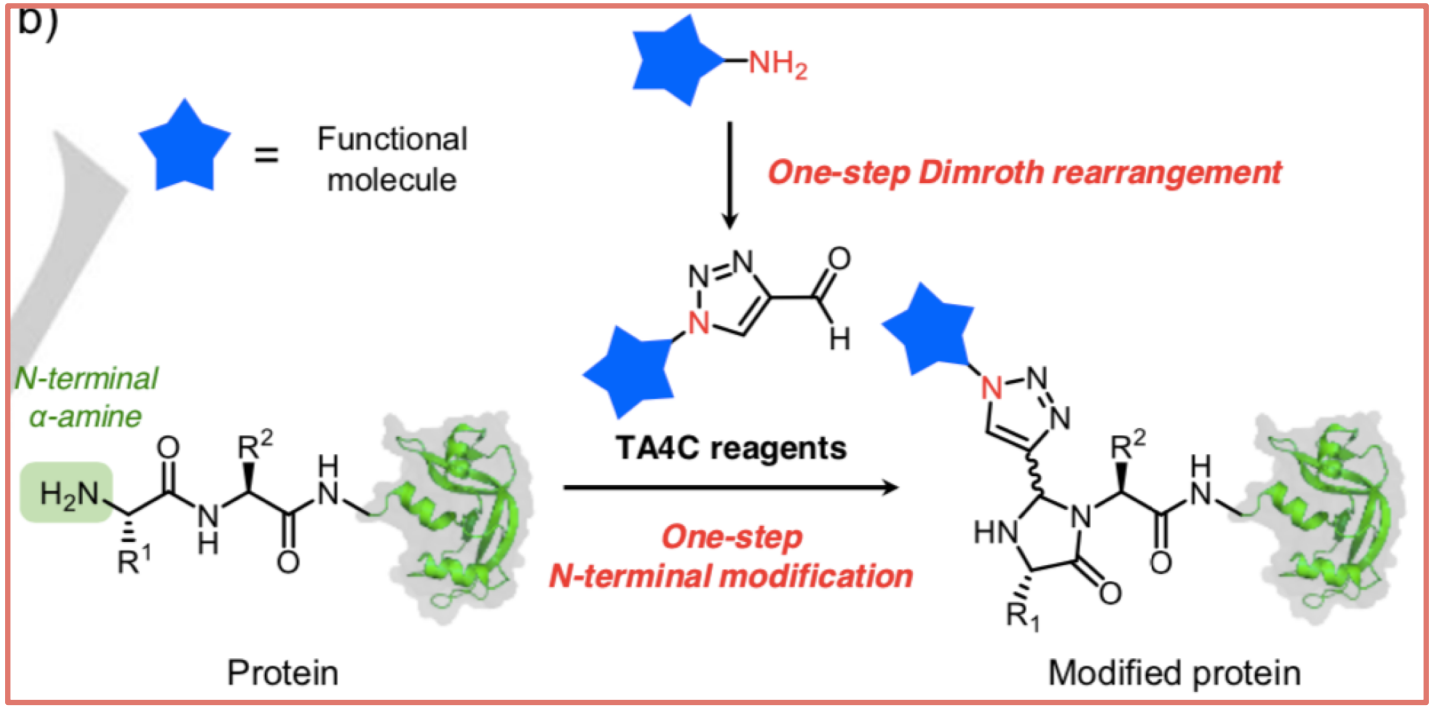
Summary of *N*-terminus selective imidazoline formation

2PCA



(MacDonald JI *et al*, 2015)

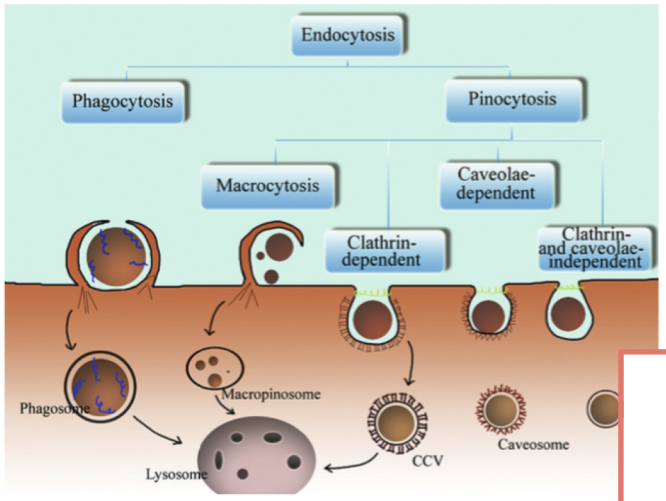
TA4C



(Onoda *et al*, 2019)

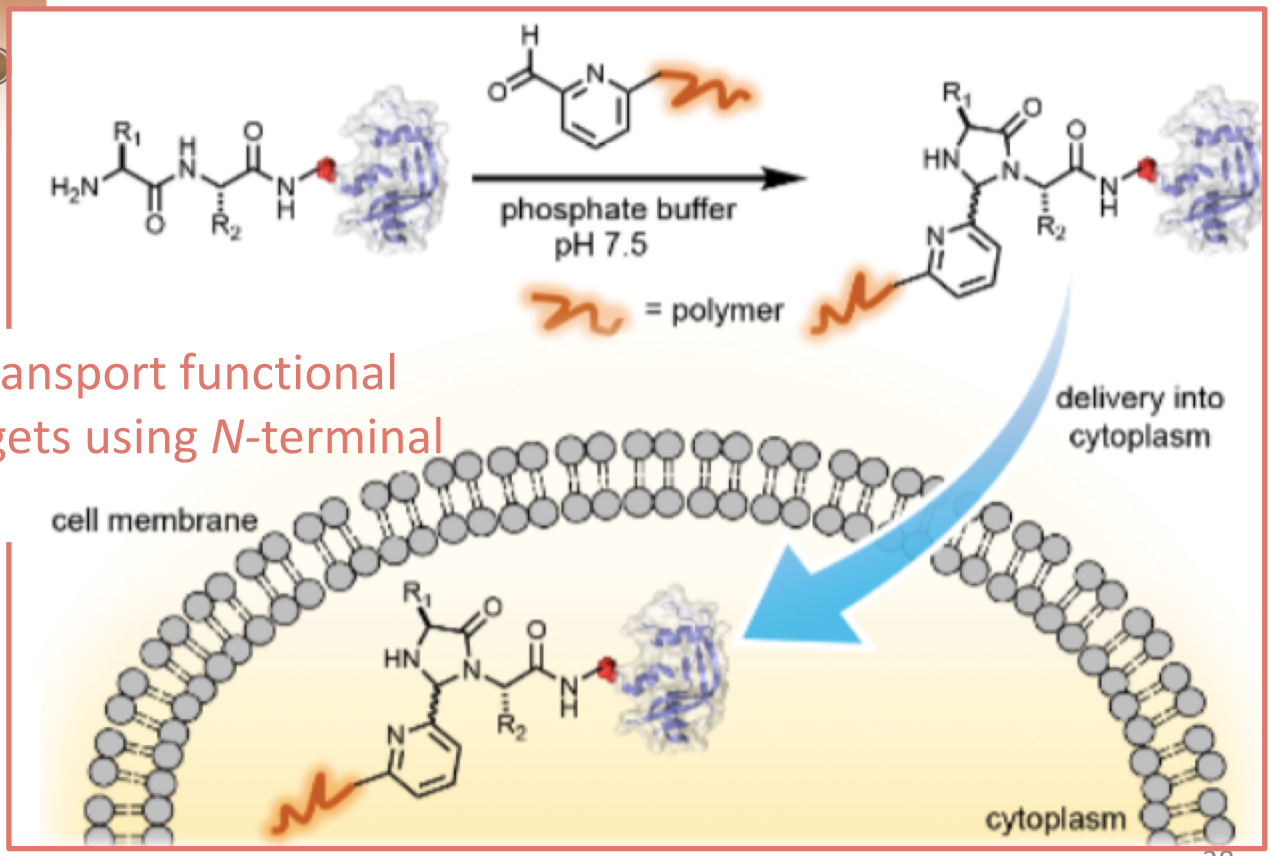
Contents

Cytosolic Delivery of Proteins



Most proteins taken up through the endocytosis pathway are destroyed in lysosomes.

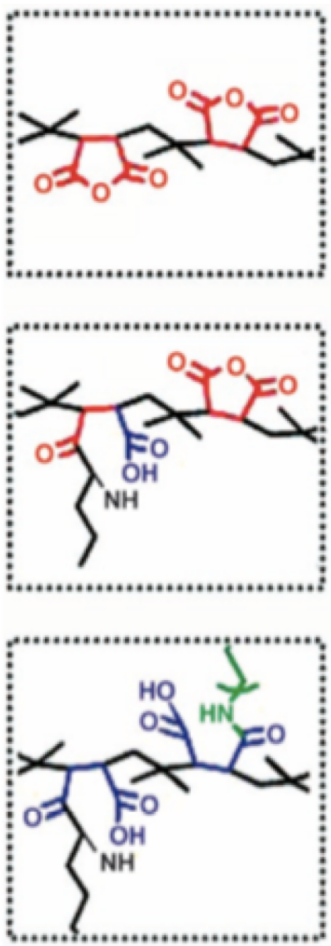
Kou, L., Sun, J., Zhai, Y., & He, Z. (2013). *Asian Journal of Pharmaceutical Sciences, 8*(1), 1-10.

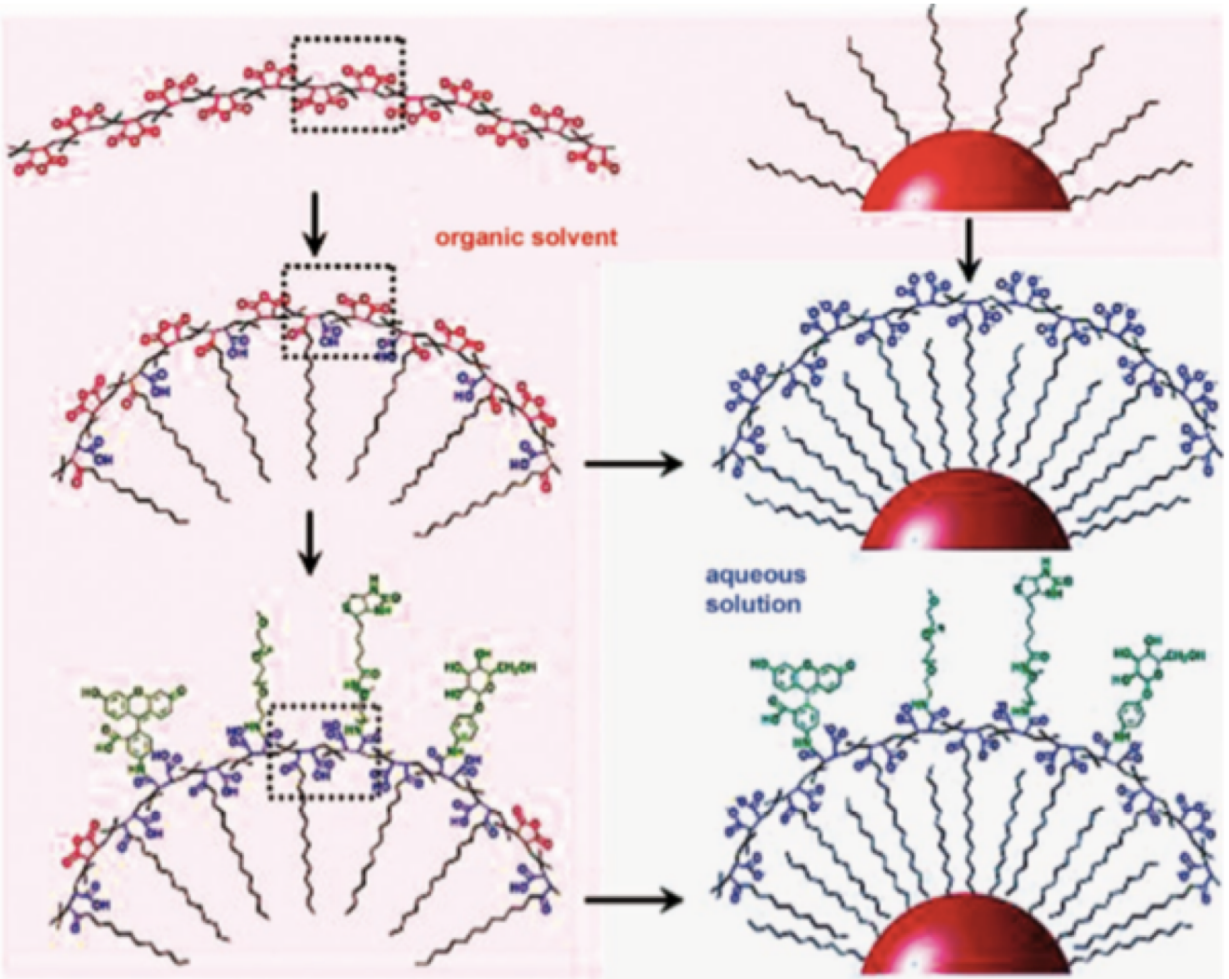


Delivery strategy that can transport functional proteins to intercellular targets using *N*-terminal modification with 2PCA.

Sangsuwan, R., Tachachartvanich, P., & Francis, M. B. (2019). *JACS, 141*(6), 2376-2383.

Amphiphilic polymer





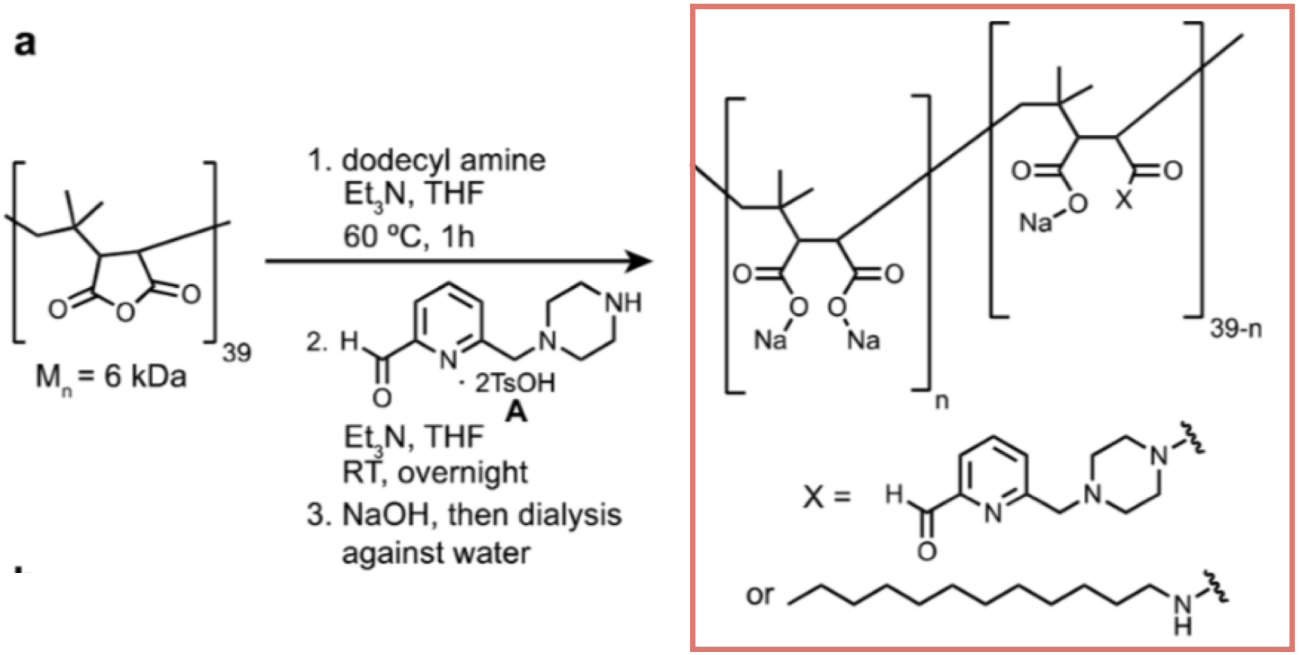
Lin, C. J., Sperling, R. A., Li, J. K., Yang, T., Li, P., Zanella, M., . . . Parak, W. J. (2008). *Small, 4*(3), 334-341.

Polymer backbone : poly(isobutylene-*alt*-maleic anhydride) (→hydrophilicity)

→ Alkylamine chains are linked by the direct amidation (→hydrophobicity)

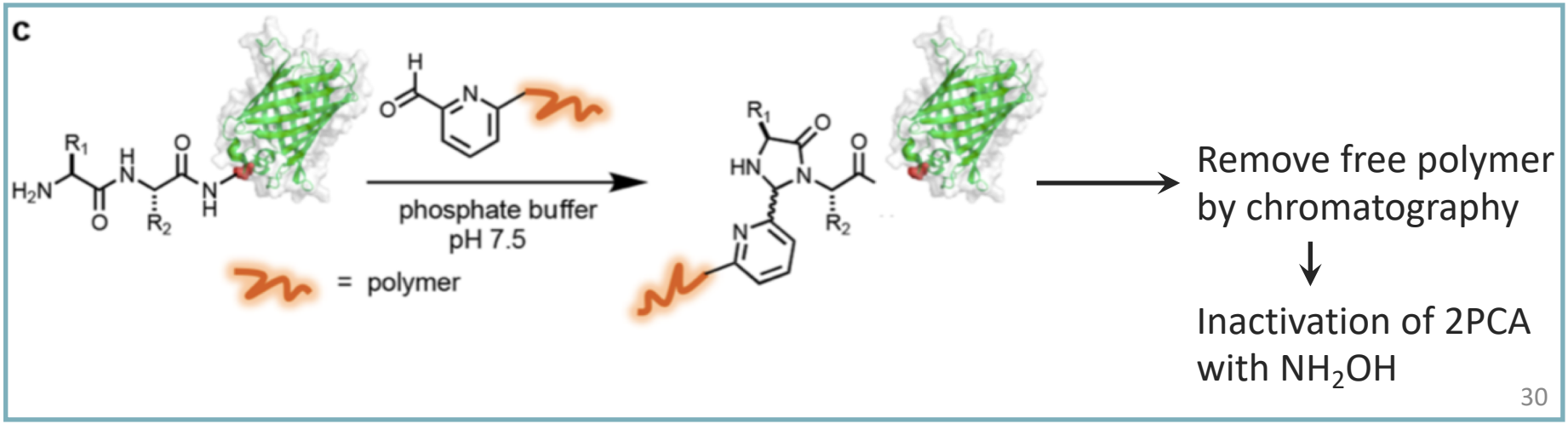
→ Functional groups with an amino terminal group are also linked to the polymer

2PCA containing polymer

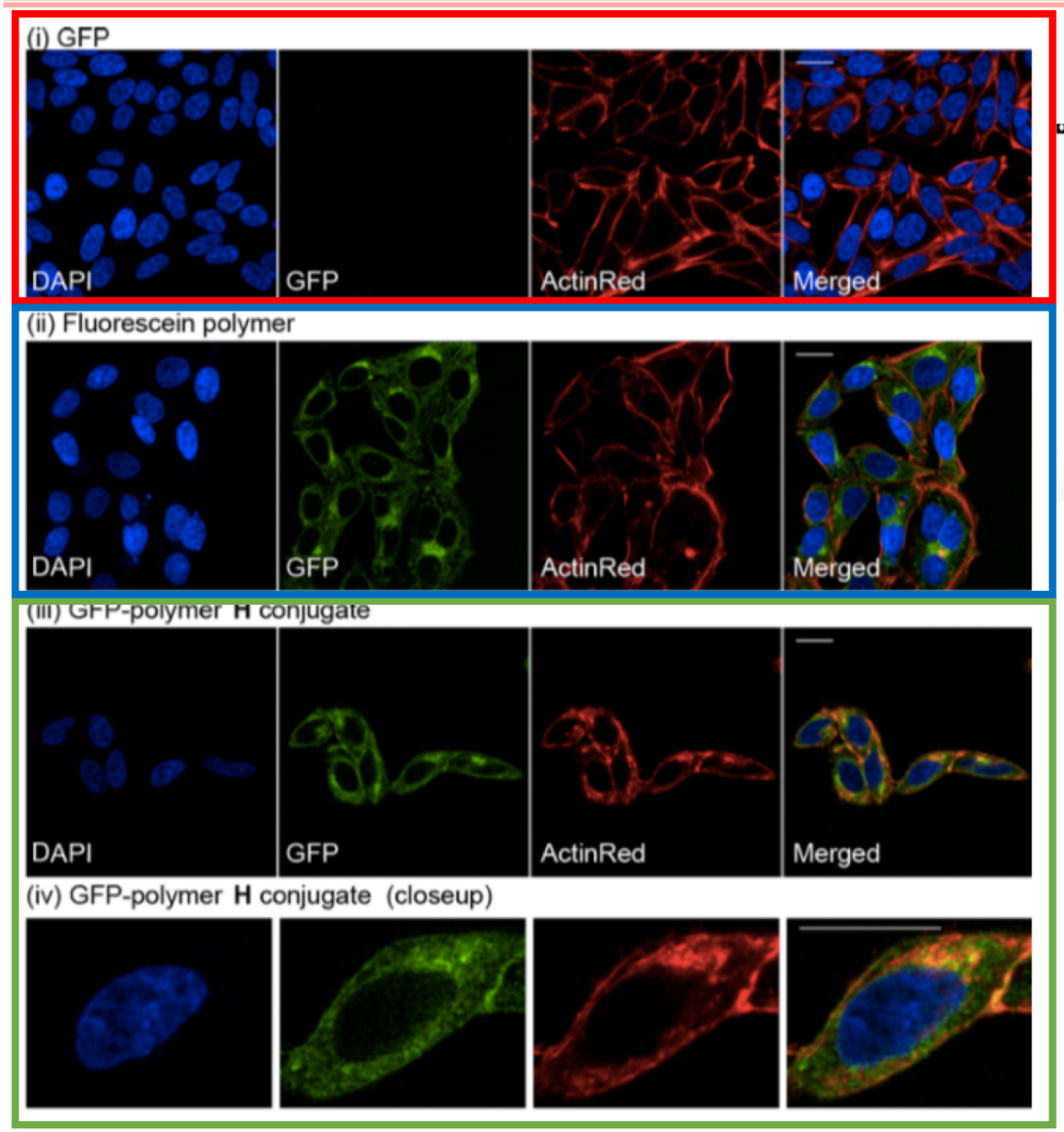


Protein-polymer conjugation

Sangsuwan, R., Tachachartvanich, P., & Francis, M. B. (2019). *JACS, 141*(6), 2376-2383.



GFP-labeled polymers in living cell



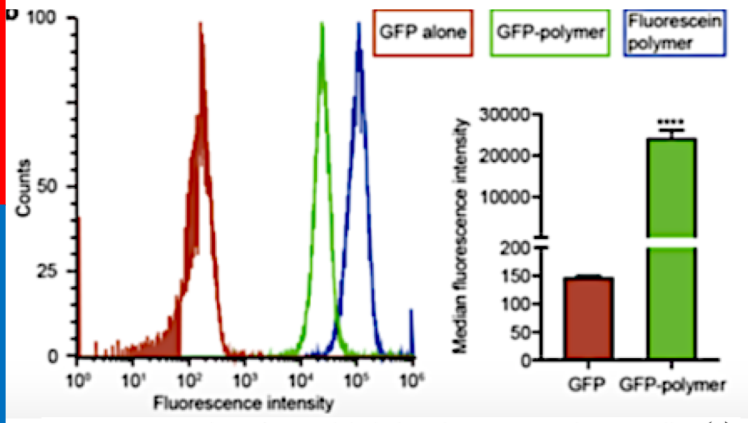


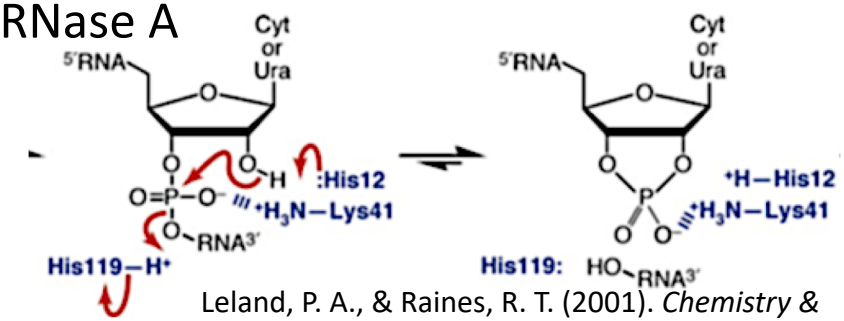
Figure 4. Uptake of GFP-labeled polymers into living cells. (a) Confocal microscopy images are shown for HeLa cells after exposure to GFP–polymer conjugates. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. Scale bars represent 20 μm. (b) HeLa cells were treated with GFP alone (10 μM), GFP conjugated to polymer **H** (10 μM), or fluorescein-labeled polymer (5 μM) for 2 h at 37 °C. Flow cytometry data are shown, along with the medians of the fluorescence intensity values. Error bars represent the standard errors resulting from three independent experiments. ****p < 0.0001 indicates a significant difference between GFP–polymer-treated cells and unmodified GFP-treated cells.

✓ Polymer conjugated protein is efficiently delivered into the living cell.

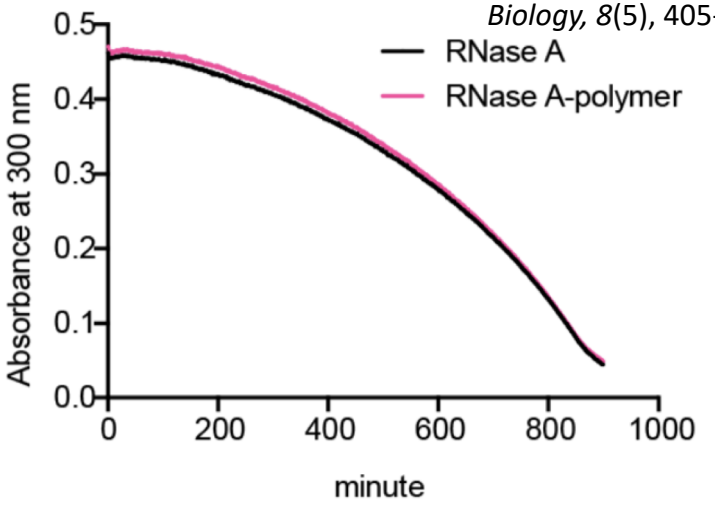
(Sangsuwan, R. *et al.* (2019))

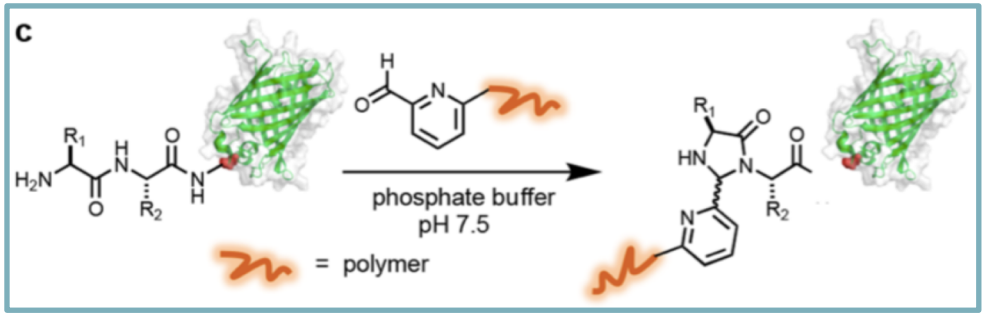
Delivery of RNase A as a Cytotoxic Cargo Protein

RNase A



Leland, P. A., & Raines, R. T. (2001). *Chemistry & Biology, 8*(5), 405-413.





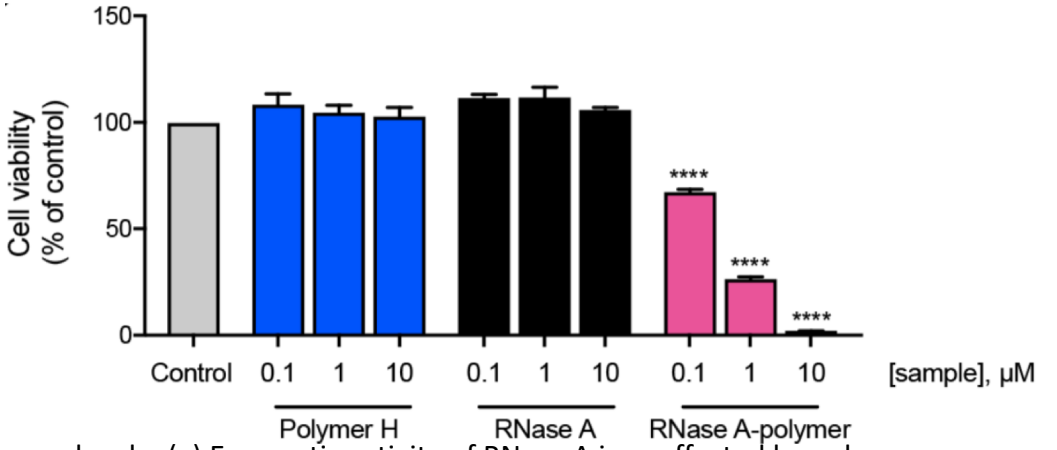


Figure 6. Delivery of RNase A to cells as a model cytotoxic cargo molecule. (a) Enzymatic activity of RNase A is unaffected by polymer attachment. Absorbance at 300 nm was recorded at 1 min intervals. (b) HeLa cells were incubated alone (control), with polymer H, or with the conjugate of RNase A to polymer H for 6 h at 37 ° C. Viability was measured using an MTT assay and is reported as the percentage of surviving cells compared to the untreated control. Error bars represent the standard error of three independent experiments. ****$p < 0.0001$ indicates a significant difference between RNase A- polymer H-treated cells and unmodified RNase A-treated cells. (Sangsuwan, R. *et al.* (2019))

- ✓ Polymer conjugation did not affect the activity of RNase A.
- ✓ RNase A-polymer exhibited cytotoxic effect.
- → Polymer conjugation did not effect the function of RNase A.

2PCA method for ADC –Synthesis of 2PCA derivative

N terminal modification with 2PCA
→Site specific modification of antibody
with high stability and selectivity under mild condition

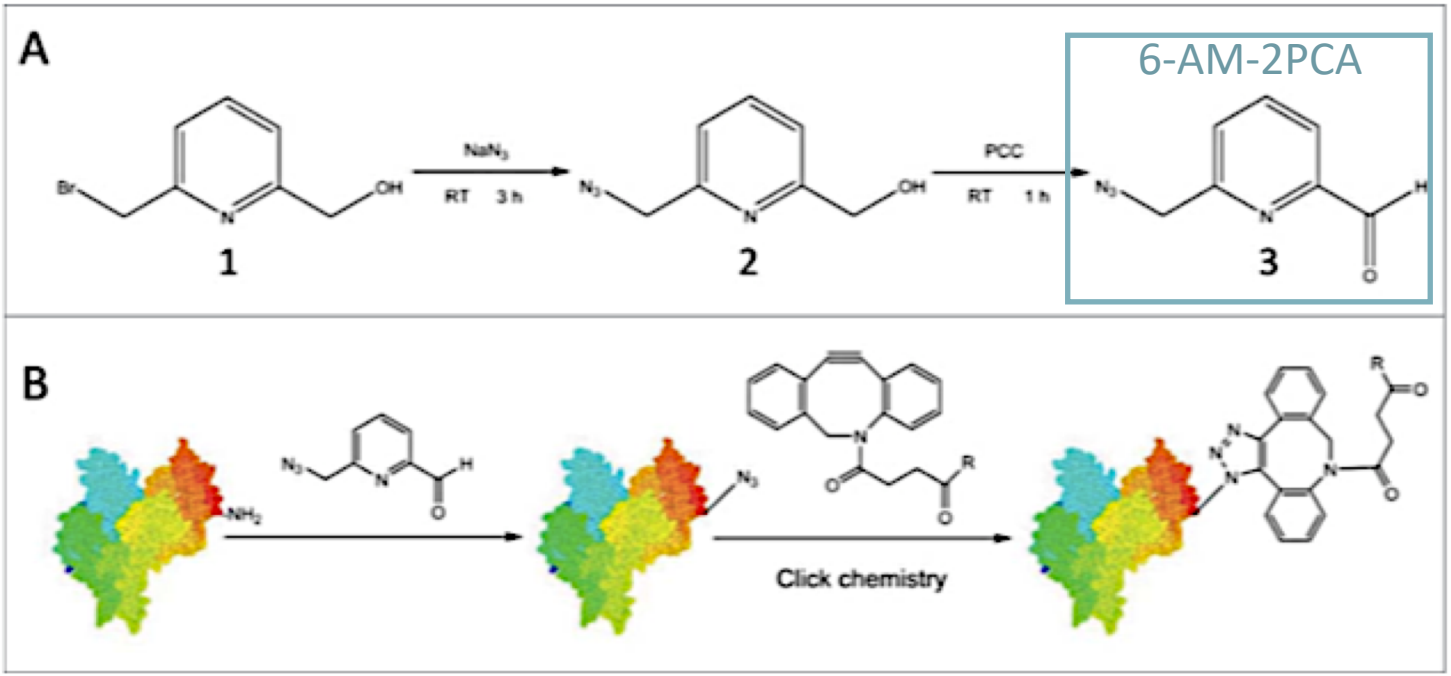


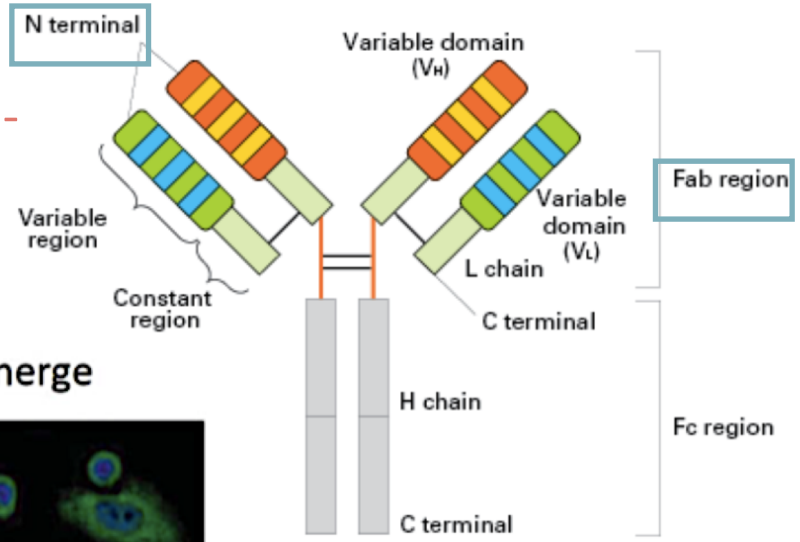
Figure 1. Synthesis of 6-AM-2-PCA and its application for N-terminal modification of peptides or proteins. (A) 6-AM-2-PCA is synthesized through azidization and subsequent oxidation of 6-(bromomethyl)-2-pyridinemethanol (1) to yield an azido intermediate (2) and ultimately, pyridinecarboxyaldehyde (3). (B) N-terminal modification of a protein with 6-AM-2-PCA results in an azide appended to the protein that orthogonally reacts with a DBCO derivative through the aldehyde-amine reaction.

Li, D., Han, B., Wei, R., Yao, G., Chen, Z., Liu, J., . . . Zhao, Q. (2018).. *mAbs, 10*(5), 712-719.

6-AM-2PCA : Azide derivative of 2PCA → click reaction

2PCA method for ADC –modification of antibody

- ✓ Fluorescence-coupled anti HER2 Fab via 6-AM-2PCA was distributed at plasma membranes.



https://www.takarabio.com/learning-centers/cdna-synthesis/cloning-antibody-variable-regions

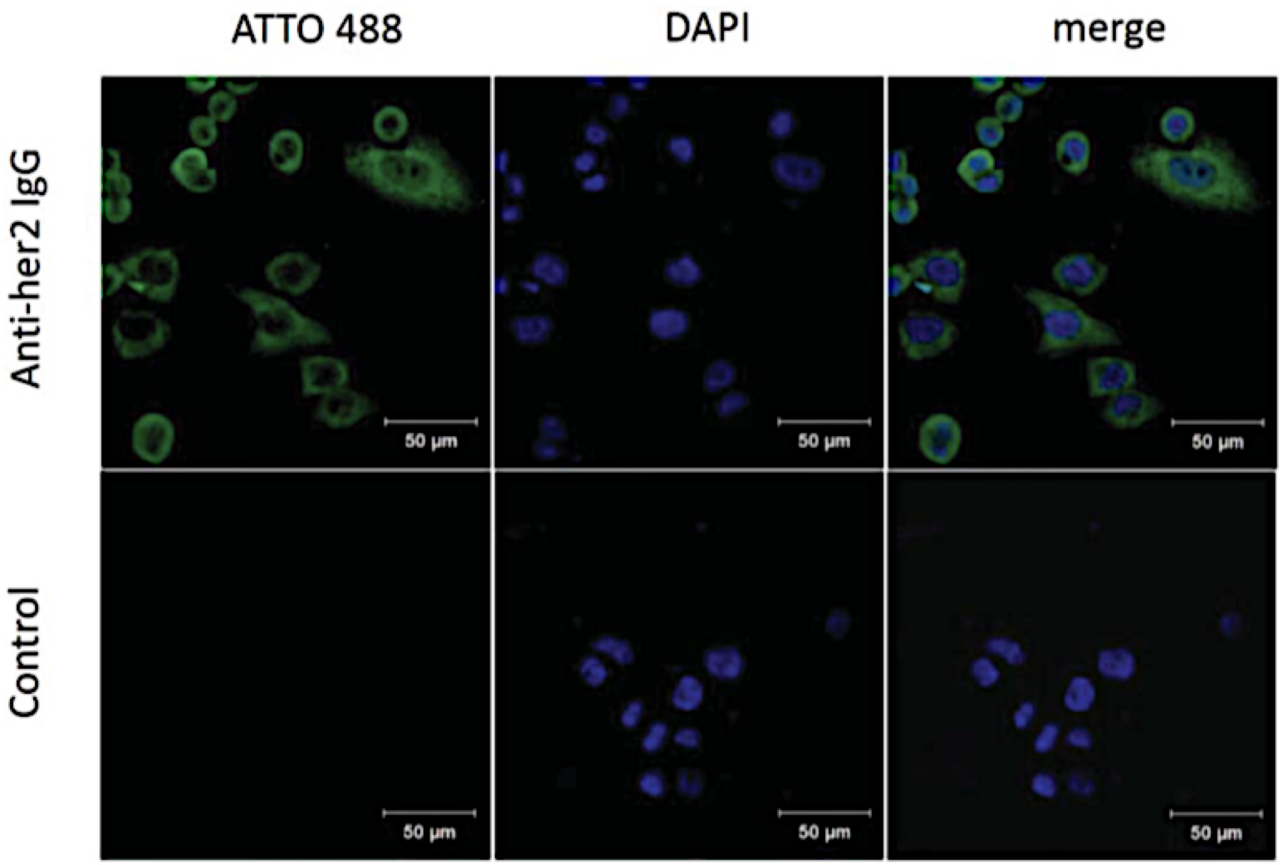


Figure 6. Immunostaining of SK-BR-3 breast cancer cells by 6-AM-2-PCA-conjugated anti-Her2 antibodies. (A) Three panels (left to right) represent fluorescence of SK-BR- 3 cells stained with modified anti-Her2 Fab followed by DBCO-ATTO 488, nuclei stained with DPAI, and the mergedimage. (B) Three panels (left to right) represent fluores-cence of SK-BR-3 cells stained with DBCO-ATTO 488 alone, nuclei stained with DPAI, and the merged image.

(Li, D. *et al.*, (2018))

Contents

Summary

- ✓ *N*-terminal residues are easily accessible, high selective modification target with little disruption to protein function.

- ✓ Imidazoline formation by 2PCA or TA4C is the favorable for *N*-terminal modification with high generality and selectivity under mild condition.

- ✓ *N*-terminal selective imidazolidinone formation would be applied to therapeutic applications including material delivery and ADC.

Appendix

Endogenous *N*-terminus modification

Table 1. Protein N-terminal modifications specified by the responsible enzymes and their substrate specificity

Nt-modification	Enzyme	Protein[b]	Alternative name	UniProt	Substrate specificity[c]
iMet excision	MetAP	MetAP1	MAP1A, Peptidase M 1	P53582	Met-Ala-, Met-Cys-, Met-Gly-, Met-Pro-, Met-Ser-, Met-Thr-, Met-Val-
		MetAP2	MAP2, MNPEP, P67EIF2	P50579	
		ND			Met-Asp-, Met-Glu-
Acetylation Propionylation[a]	NatA	$Naa10^{cat}$	ARD1, ARD1A, TE2	P41227	Ala-, Cys-, Gly-, Ser-, Thr-, Val-, Asp[d]-, Glu[d]-
		$Naa15^{aux}$	NAT1, GA19, NARG1, NATH, TBN	Q9BXJ9	
	NatB	$Naa20^{cat}$	NAT3, NAT5	P61599	Met-Asn-, Met-Asp-, Met-Gln-, Met-Glu-
		$Naa25^{aux}$	MDM20, NAP1	Q14CX7	
	NatC	$Naa30^{cat}$	MAK3, NAT12	Q1473	Met-Ile-, Met-Leu-, Met-Phe-, Met-Trp-
		$Naa35^{aux}$	MAK10, EGAP,	Q5VZE5	
		$Naa38^{aux}$	MAK31, LSMD1, PFAAP2	Q9BRA0	
	NatD	$Naa40^{cat}$	NAT4, NAT11	Q86UY6	Ser-Gly-Gly-, Ser-Gly-Arg-
	NatE	$Naa50^{cat}$	MAK3, NAT13, SAN	Q9GZZ1	Met-Ala-, Met-Leu-, Met-Lys-, Met-Phe-, Met-Ser-, Met-Thr-, Met-Tyr-, Met-Val-
	NatF	$Naa60^{cat}$	NAT15, HAT4	Q9H7X0	Met-Ala-, Met-Gln-, Met-Gly-, Met-Ile-, Met-Leu-, Met-Lys-, Met-Met-, Met-Ser-, Met-Thr-, Met-Tyr-, Met-Val-
Methylation	NTMT	NTMT1	METTL11A, NRMT1, NRMT1A	Q9BV86	Ala/Pro/Ser-Pro-Lys-
		Tae1 (*S.c*)	NTM1	P38340	
		NTMT2	METTL11B, NRMT2, NTM1B	Q5VVY1	
Myristoylation	NMT	NMT1	NMT	P30419	Gly-
		NMT2		O60551	
Palmitoylation	PAT	Hhat	MART2, SKI1, Skn	Q5VTY9	Cys-
		Rasp (*D.m*)	cmn, sit, ski	Q9VZU2	
		ND			Gly-
Ubiquitylation	Ubiquitin ligase	Ube2w	UBC16, UBC-16,	Q96B02	Unstructured N-terminal backbone
		HUWE1	ARF-BP1, HectH9, LASU1, Mule, UREB1, URE-B1	Q7Z6Z7	*ND*

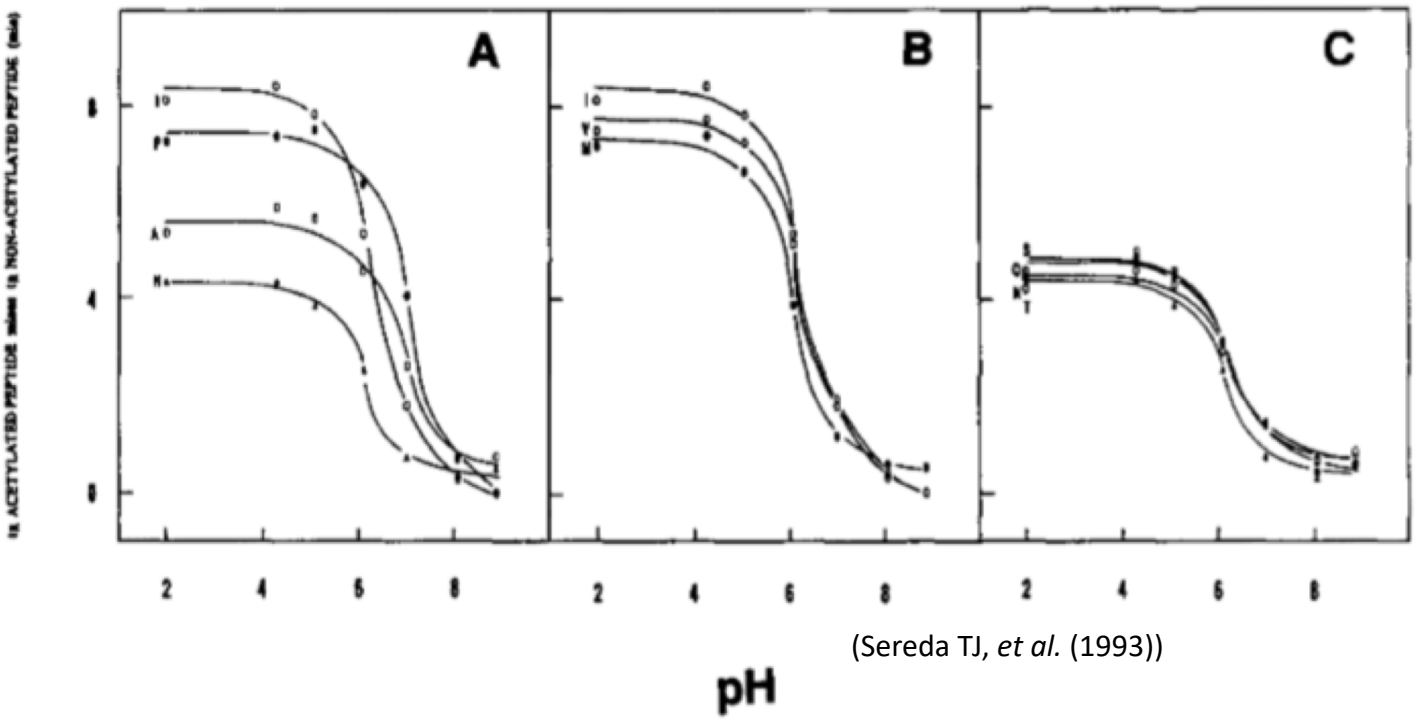
a) N-terminal propionylation is catalyzed by the same enzymes as N-terminal acetylation (NATs), and the substrate specificity is presumably shared.
b) All proteins listed are human except where the species is indicated.
c) The indicated amino acid sequences do not guarantee N-terminal modification.
d) Naa10 substrates.
ND, not determined; Aux, auxiliary subunit; Cat, catalytic subunit; *S.c*, *Saccharomyces cerevisiae*; *D.m*, *Drosophila melanogaster*.

Varland S, Osberg C, Arnesen T., *PROTEOMICS*. 2015;15(14):2385-2401.

✓ Most proteins expressed in bacteria and some protein expressed in eukaryotes including antibodies could be the target of artificial *N*-terminal modification.

pKa

$t_{R_{Ac\text{-}Xx}} - t_{R_{H\text{-}Xx}}$



(Sereda TJ, *et al.* (1993))

Fig. 5. Plot of retention time of acetylated peptide analogue minus the retention time of the non-acetylated peptide analogue *versus* pH. Column: PLRP-S (250 × 4.6 mm I.D., 5 μm, 100 Å). Mobile phase: linear A–B gradient elution (2% B/min equivalent to 1% acetonitrile/min) at a flow-rate of 1 ml/min. In the *pH 2* system, A is 20 m*M* aqueous H_3PO_4 containing 2% acetonitrile and B is 20 m*M* H_3PO_4 in acetonitrile–water (1:1); in the *pH 4–7* system, A is 20 m*M* aqueous triethylammonium phosphate (TEAP) containing 2% acetonitrile and B is 20 m*M* TEAP in acetonitrile–water (1:1). In the *pH 7–9* system, linear A–B gradient elution (2% B/min, equivalent to 1% acetonitrile/min), A is 10 m*M* aqueous $(NH_4)_2HPO_4$ containing 2% acetonitrile and B is 10 m*M* $(NH_4)_2HPO_4$ in acetonitrile–water (1:1), both eluents containing 100 m*M* sodium perchlorate. Panels A–C represent 11 examples of the peptide analogues used in this study.

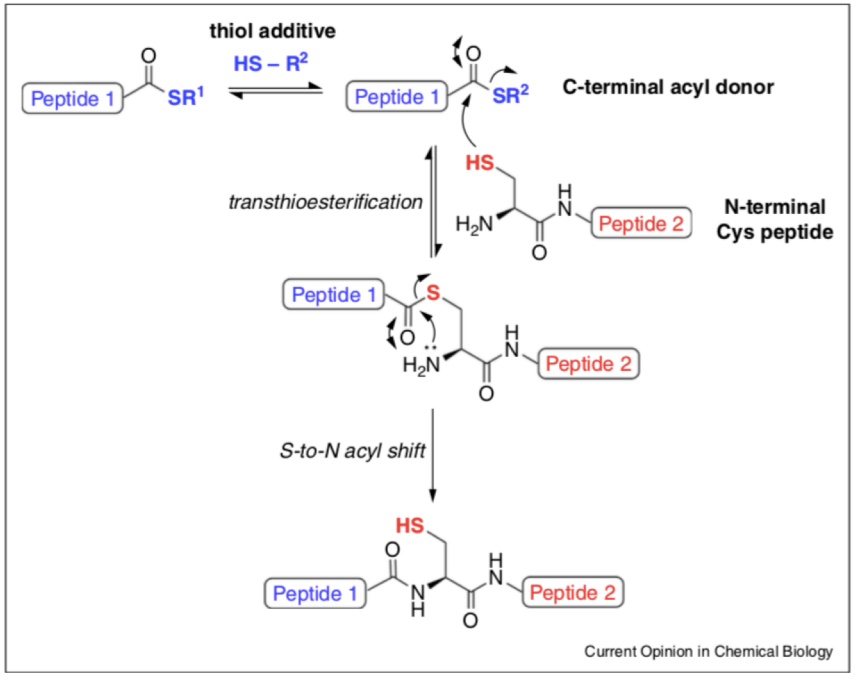
Model peptide

Ac-X-LGAKGAGVG-amide

H-X-LGAKGAGVG-amide

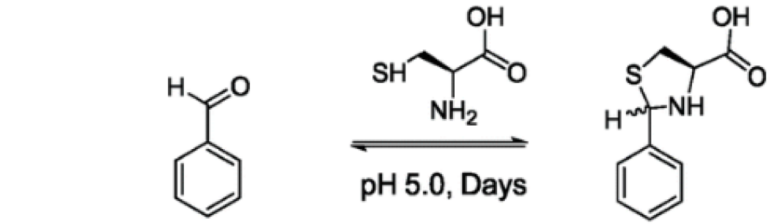
Modification of Cysteine at protein N terminus

Native chemical ligation

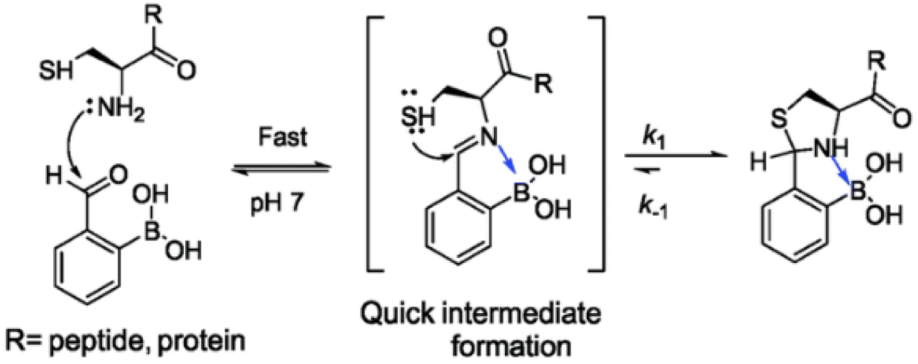


Malins, L. R., & Payne, R. J. (2014). *Current Opinion in Chemical Biology, 22*, 70-78.

Forming thiazolidine with aldehyde



This Work



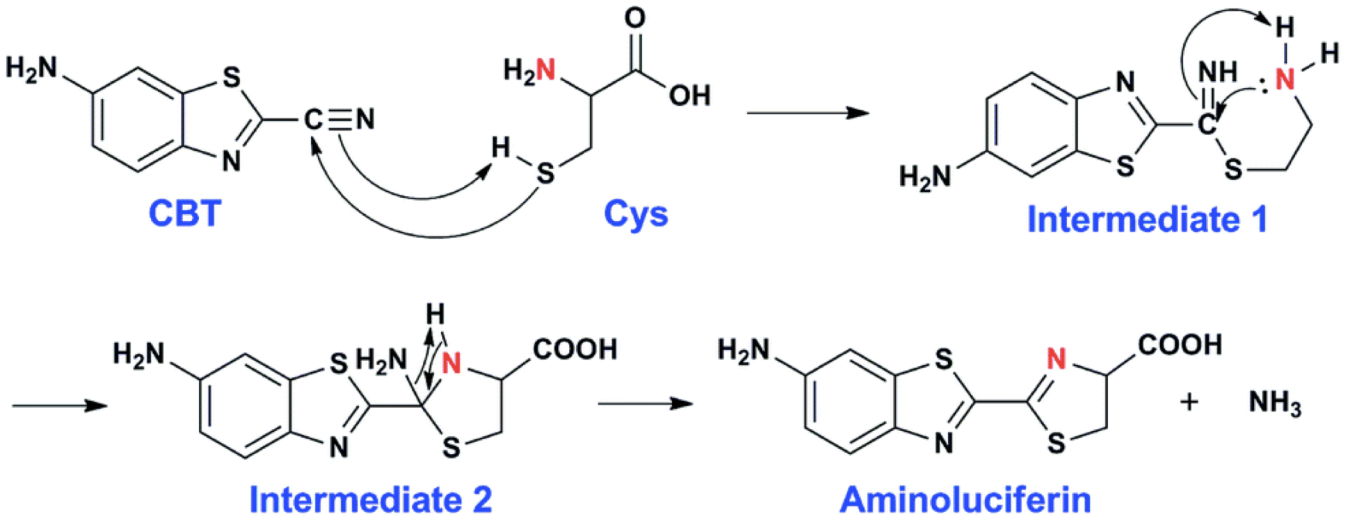
Bandyopadhyay, A.; Cambray, S.; Gao, (2016) J. *Chem. Sci.* , *7*, 4589.

- The reaction is relatively slow (48 h).
- It requires acidic condition.
- A large excess of the aldehyde is needed.

→*ortho-boronic* stabilizes the thiazolidino-boronate product through *N→B* coordination.

Modification of Cysteine at protein N terminus

Condensation with 2-cyanobenzothiazole derivatives

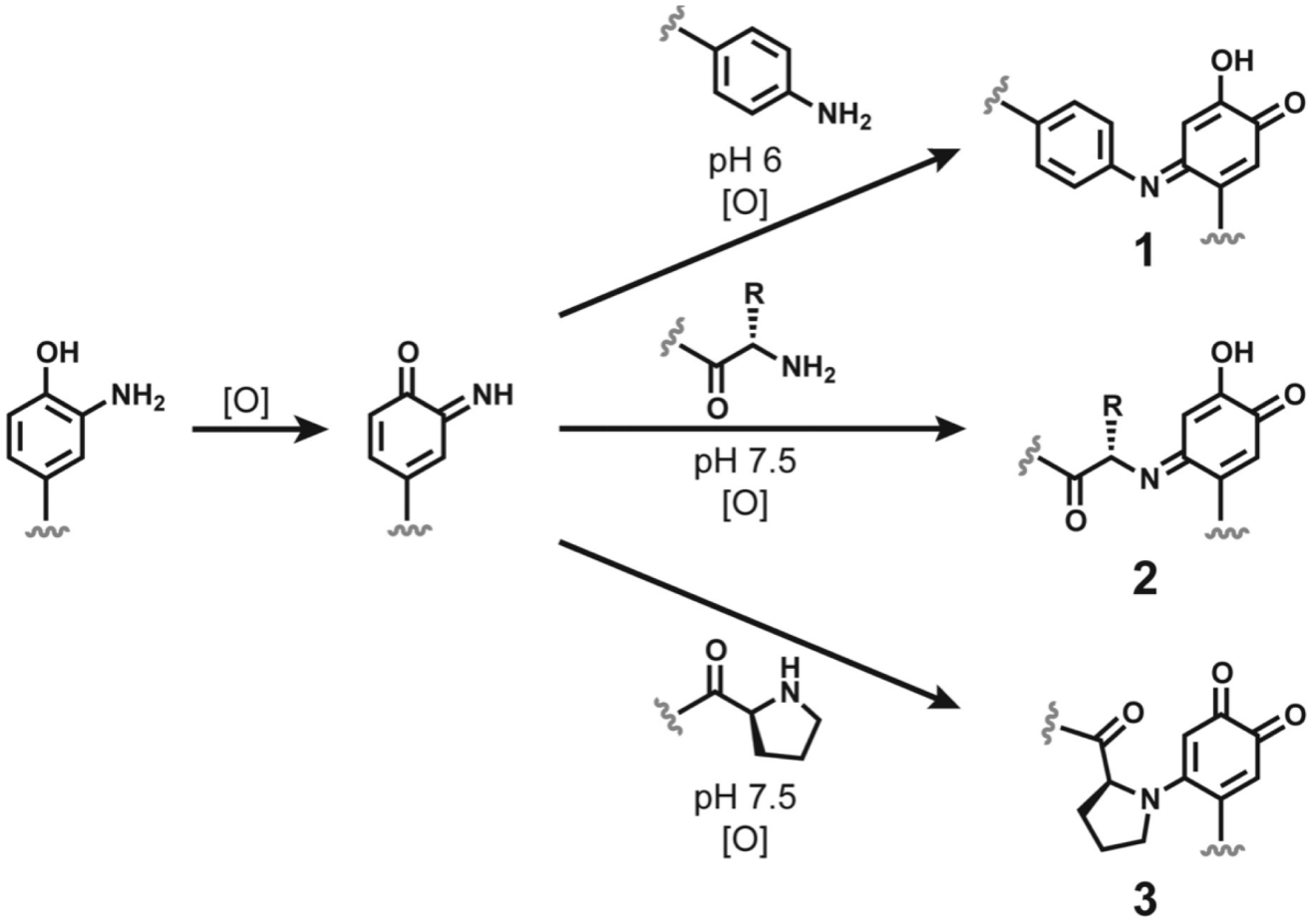


Zheng, Z., Chen, P., Li, G., Zhu, Y., Shi, Z., Luo, Y., . . . Liang, G. (2017). *Chemical Science, 8*(1), 214-222。

◯ The reaction is rapid at neutral pH and generates stable product.

→CBT reacts with $_{D}$-cysteine in the final step of $_{D}$-luciferin synthesis.

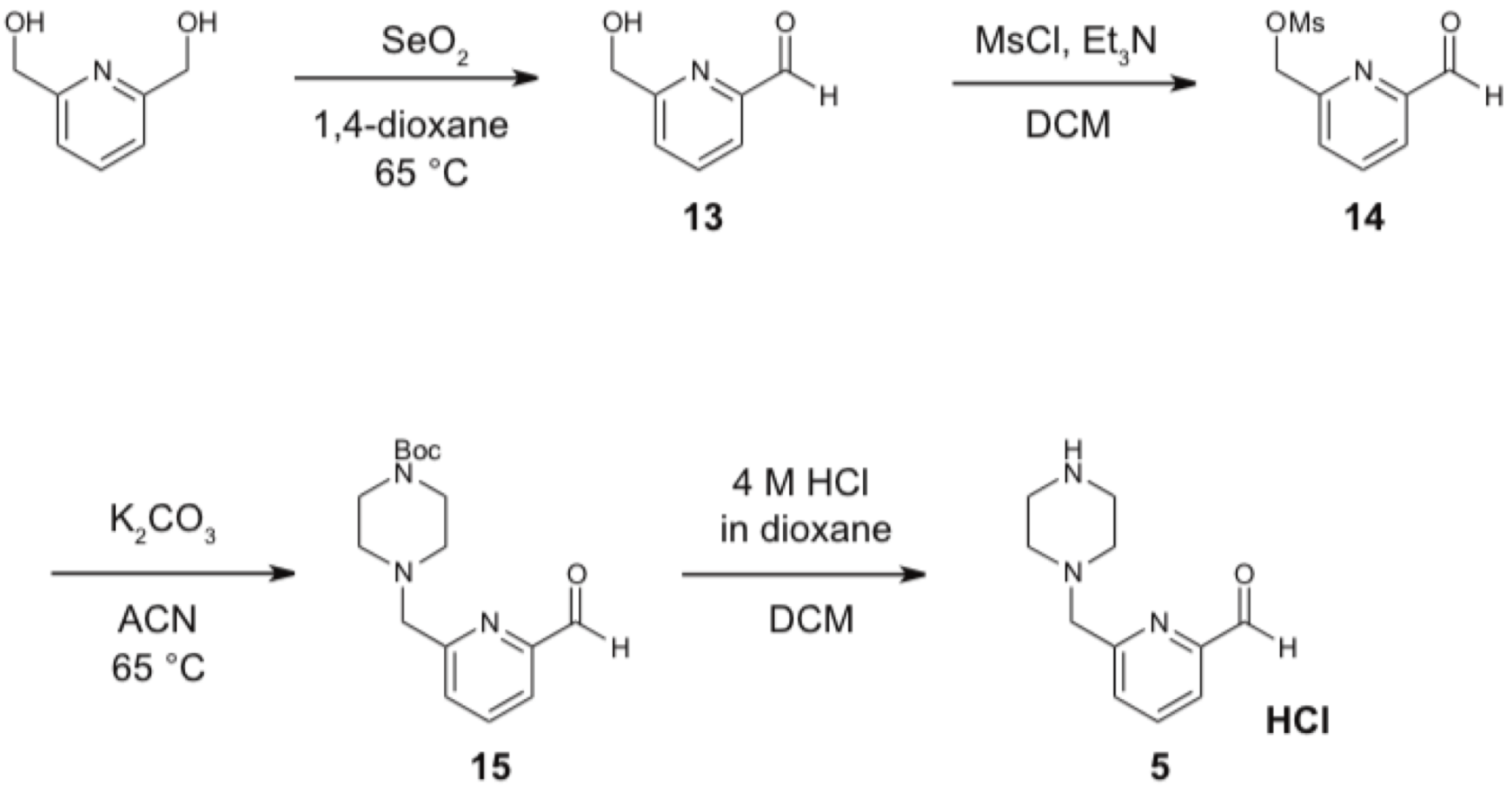
Oxidative coupling



Obermeyer, A. C., Jarman, J. B., & Francis, M. B. (2014). *JACS*, *136*(27), 9572-9579.

- Proline still stood out as the most reactive species.
- Free cysteine residues can be oxidized.

Synthesis of 2PCA intermediate



(MacDonald JI *et al*, 2015)

Protein scope for N-terminal biotunylation

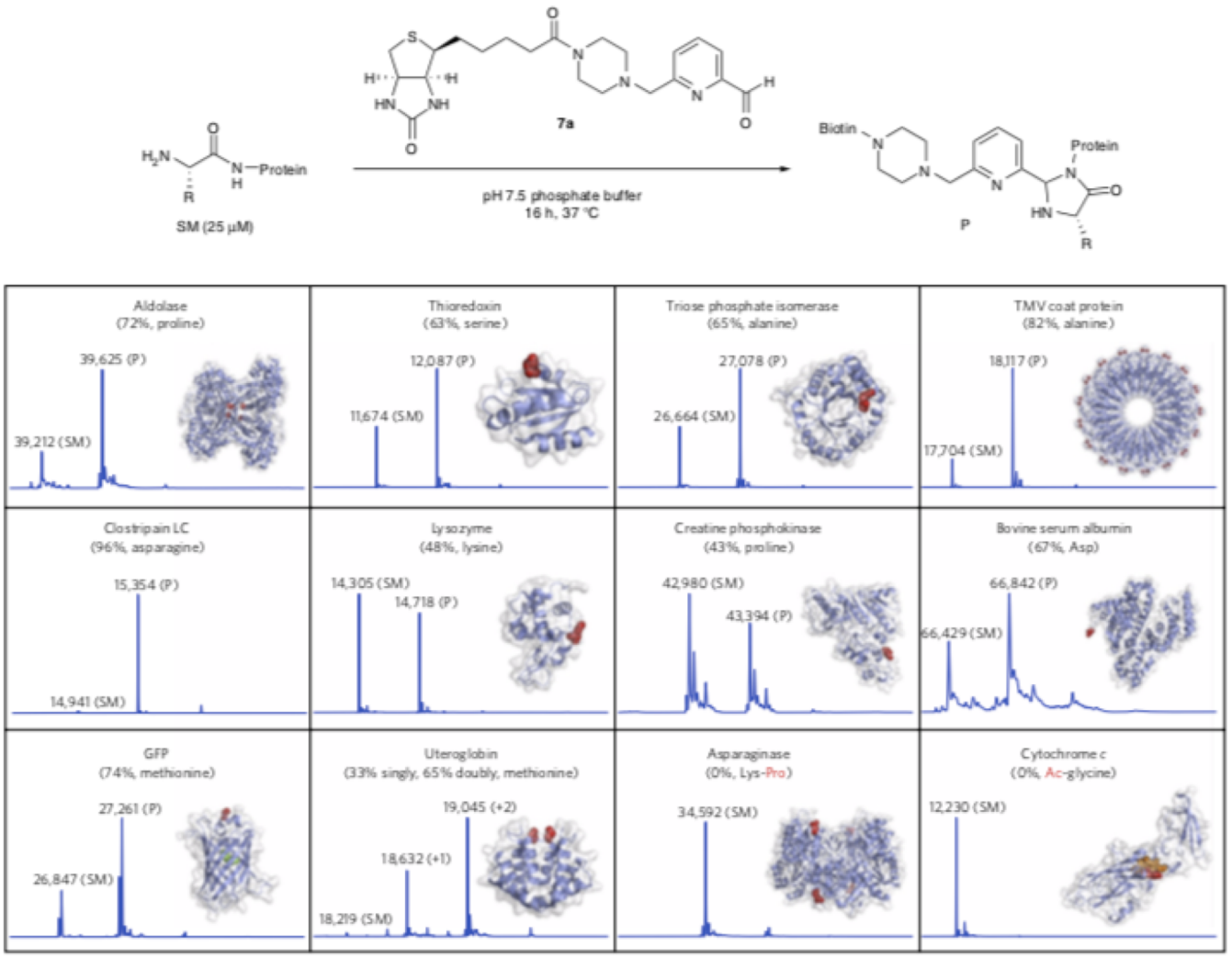
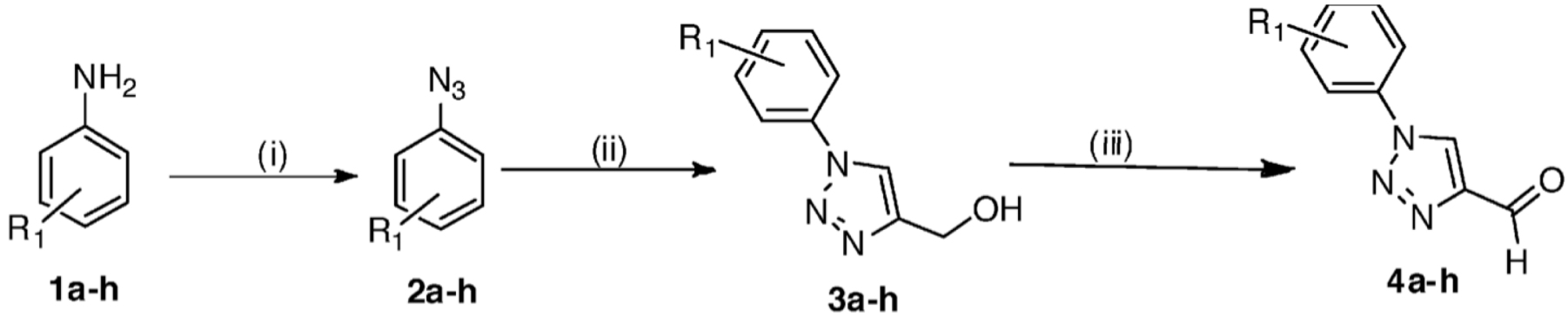


Figure 3 | Site-specific attachment of 2PCA-biotin reagent 7a to ten protein substrates, as characterized using ESI-TOF MS. For each entry, the molecular weight and the N-terminal amino acids are specified. For clarity, all starting proteins are labeled 'SM', and species corresponding to the correct product mass are labeled 'P'. Uteroglobin is modified up to two times because the protein is a disulfide dimer and has two N termini. Asparaginase is not modified because it has a proline residue in position 2, and cytochrome *c* is acylated at the N terminus. These latter examples serve as negative controls, underscoring the fact that protein side chains do not participate in this chemistry. Conditions: 25 μM protein, 10 mM **7a**, 50 mM phosphate buffer at pH 7.5 and 37 °C for 16 h.

(MacDonald JI *et al*, 2015)

Synthesis of TA4C precursor

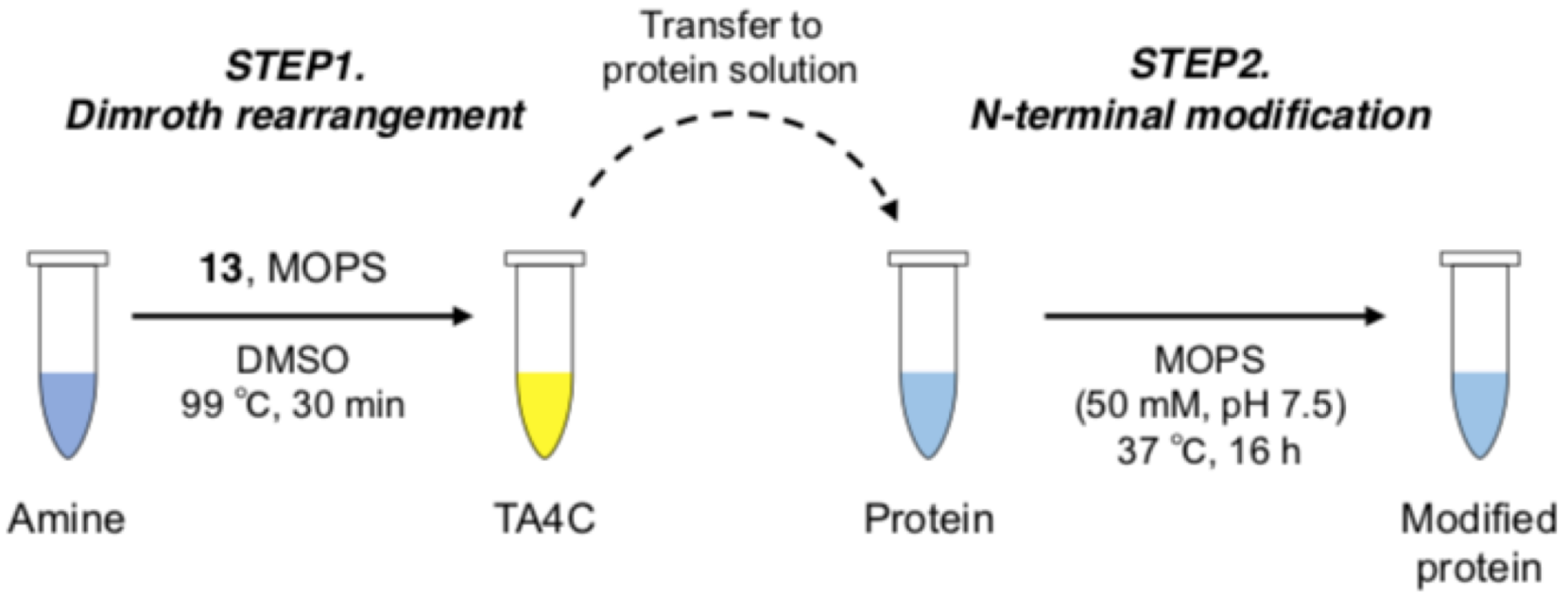


R_1 = H, 2-Cl, 4-Cl, 2-OMe, 4-OMe, 4-OH, 3-CF_3, 4-NO_2

Reagents and conditions: (i) $NaNO_2$, HCl 10%; NaN_3, 2-4 hr, rt;
(ii) propargyl alcohol, $CuSO_4.5H_2O$, sodium ascorbate. H_2O, DMF, 3 hr, rt;
(iii) IBX/DMSO, 3 hr, rt.

Goud, G. L., Ramesh, S., Ashok, D., Reddy, V. P., Yogeeswari, P., Sriram, D., ... & Manga, V. (2017), *MedChemComm*, *8*(3), 559-570.
(Search in SciFinder)

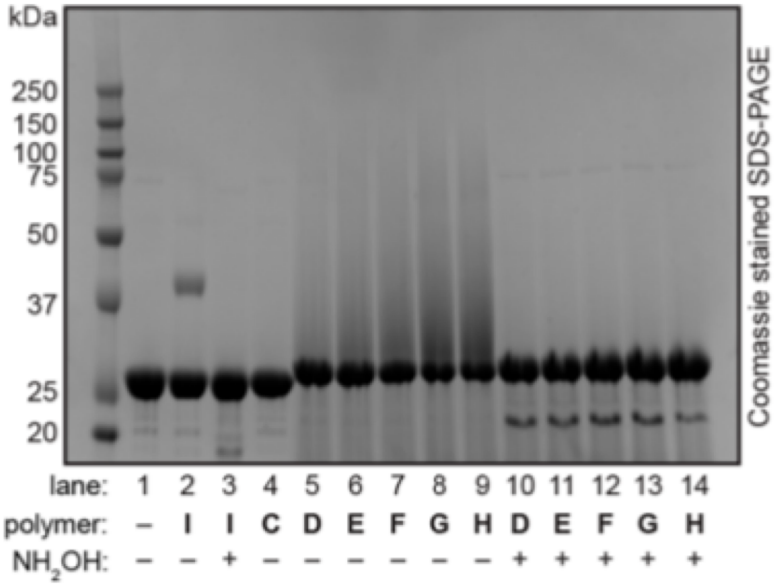
Two step modification with TA4C



An amine precursor in DMSO (200 mM, 20 µL, final concentration: 100 mM) and 13 in DMSO (200 mM, 20 µL, final concentration: 100 mM) were mixed in a tube (1 mL). To the solution was added MOPS in H2O (400 mM, 1 µL, final concentration: 10 mM), and the mixture was incubated at 99 ° C for 30 min. The Dimroth rearrangement reaction was assessed by 1H NMR measurement (400 MHz, DMSO-d6) (Figure S15~S18).
RNase in H2O (1 mM, 2.5 µL, final concentration: 50 µM) was diluted with potassium phosphate buffer (10 mM, pH 7.5, 42.5 µL). To the resulting solution, a crude solution of TA4C 1, 10, 14 or 15 in DMSO (5.0 µL) was added, and the mixture was incubated at 37 ° C for 16 h. The protein was then purified several times by centrifugal filtration in an Amicon Ultra-0.5 centrifugal filter device with a 10-kDa molecular weight cutoff (Millipore). Modification of proteins was analyzed by LC-MS.

Polymer variants

polymer	equiv. of **A**	actual **A**/polymer
C	0	0
D	1	2
E	5	4
F	10	5
G	20	12
H	40	16



- ✓ Polymer H showed the greatest degree of conjugation (~68 %) and greater cellar uptake.

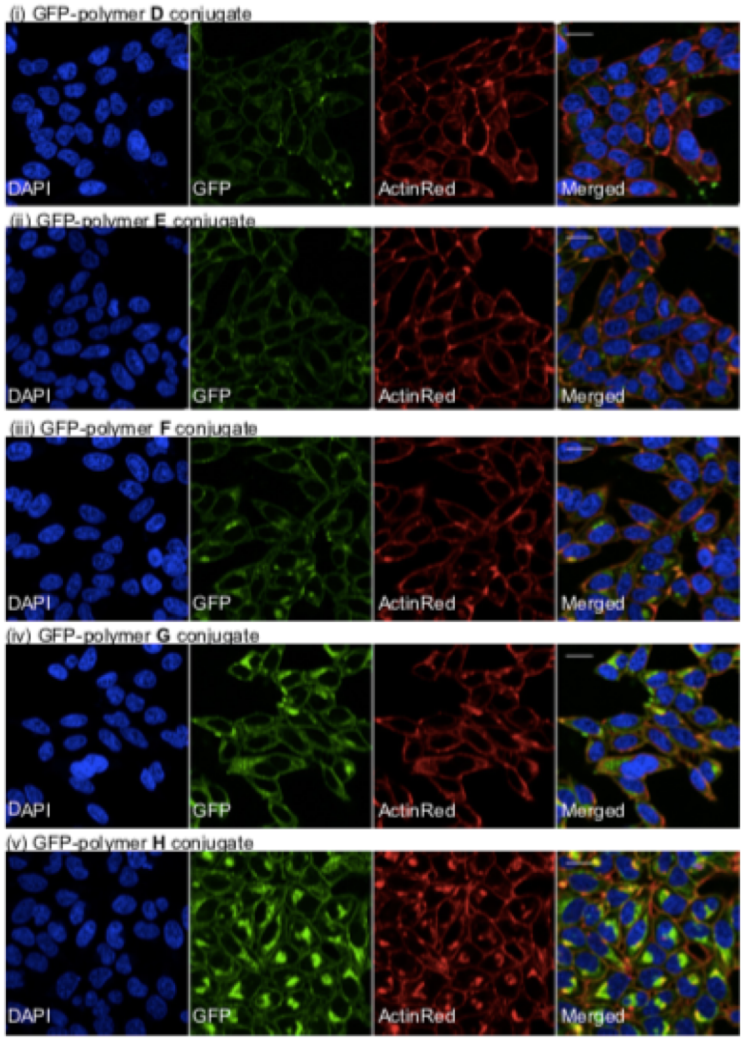
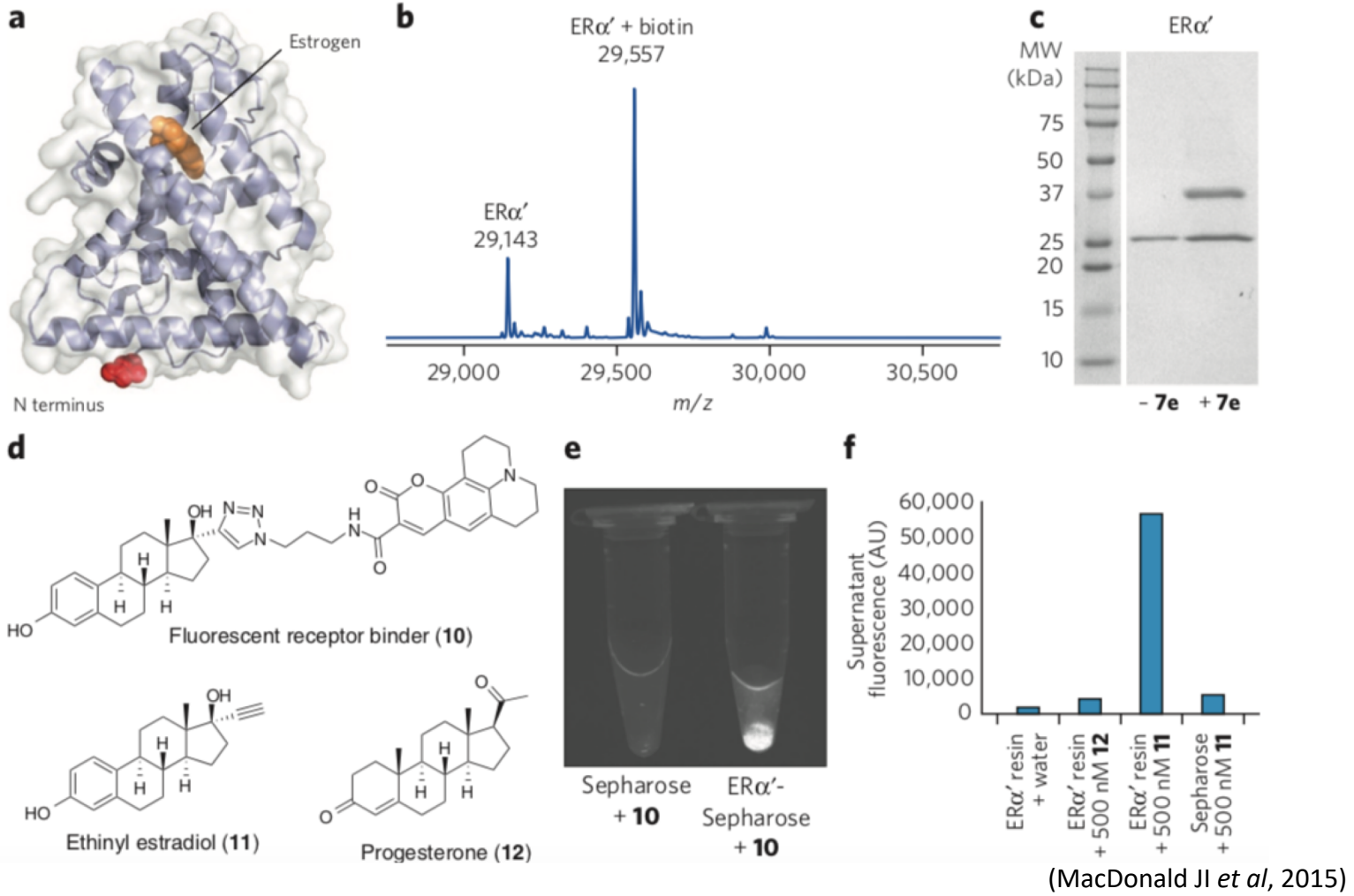


Figure S3. The confocal microscopy study of the conjugates of GFP to polymers **D-H** in HeLa cells. Nuclear and cytoplasmic stains were performed on fixed cells using DAPI and ActinRed, respectively. Scale bars represent 20 μm. Cells were treated with the conjugates of GFP to polymers **D-H** (10 μM) for 2 h at 37 ℃.

Polymer immobilization of the ligand-binding domain of Erα' (application)



(MacDonald JI *et al*, 2015)

2PCA incorporates high-affinity binding proteins for endocrine-disrupting compounds into polymer matrix that can be easily detected.

→Detection of those compounds in drinking water.

Capture and Recycling of SrtA (application)

Methods for the selective capture and subsequent recycling of enzyme were in high demand.

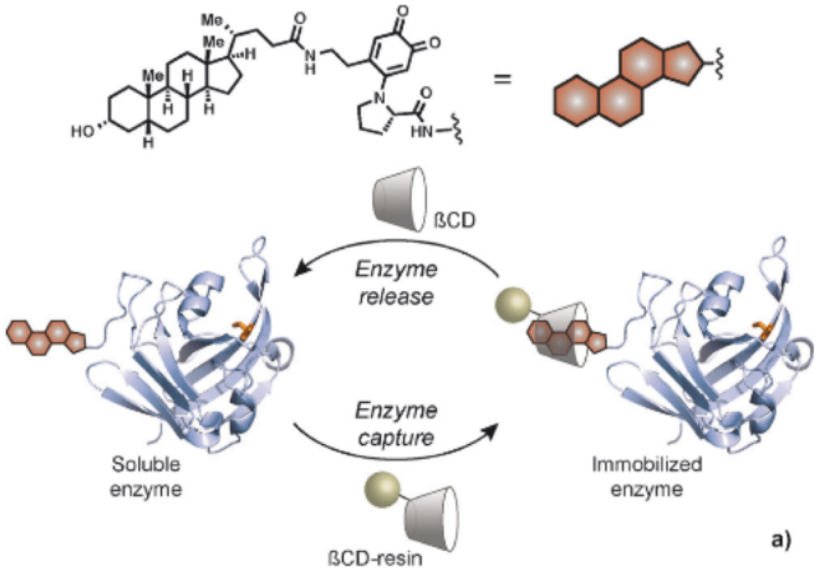
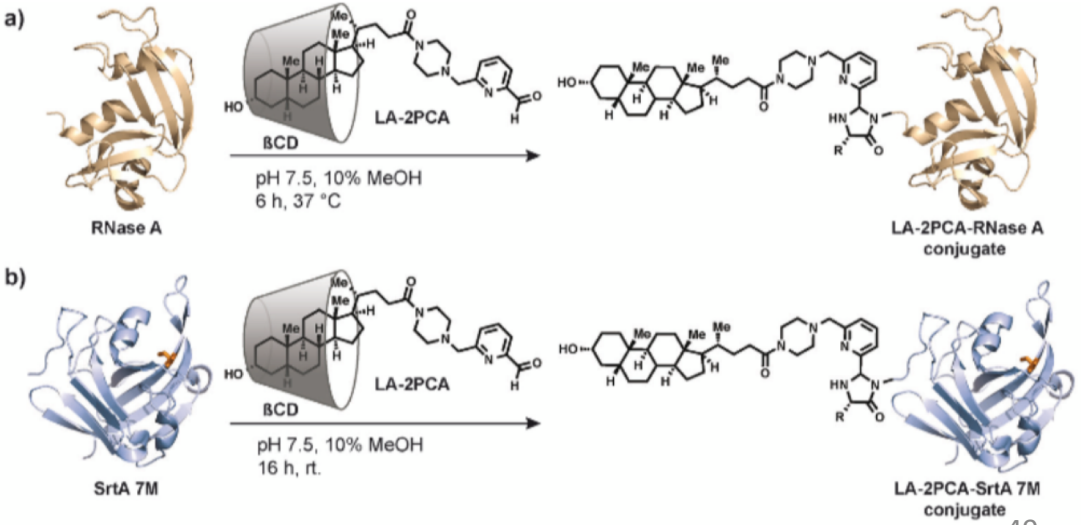


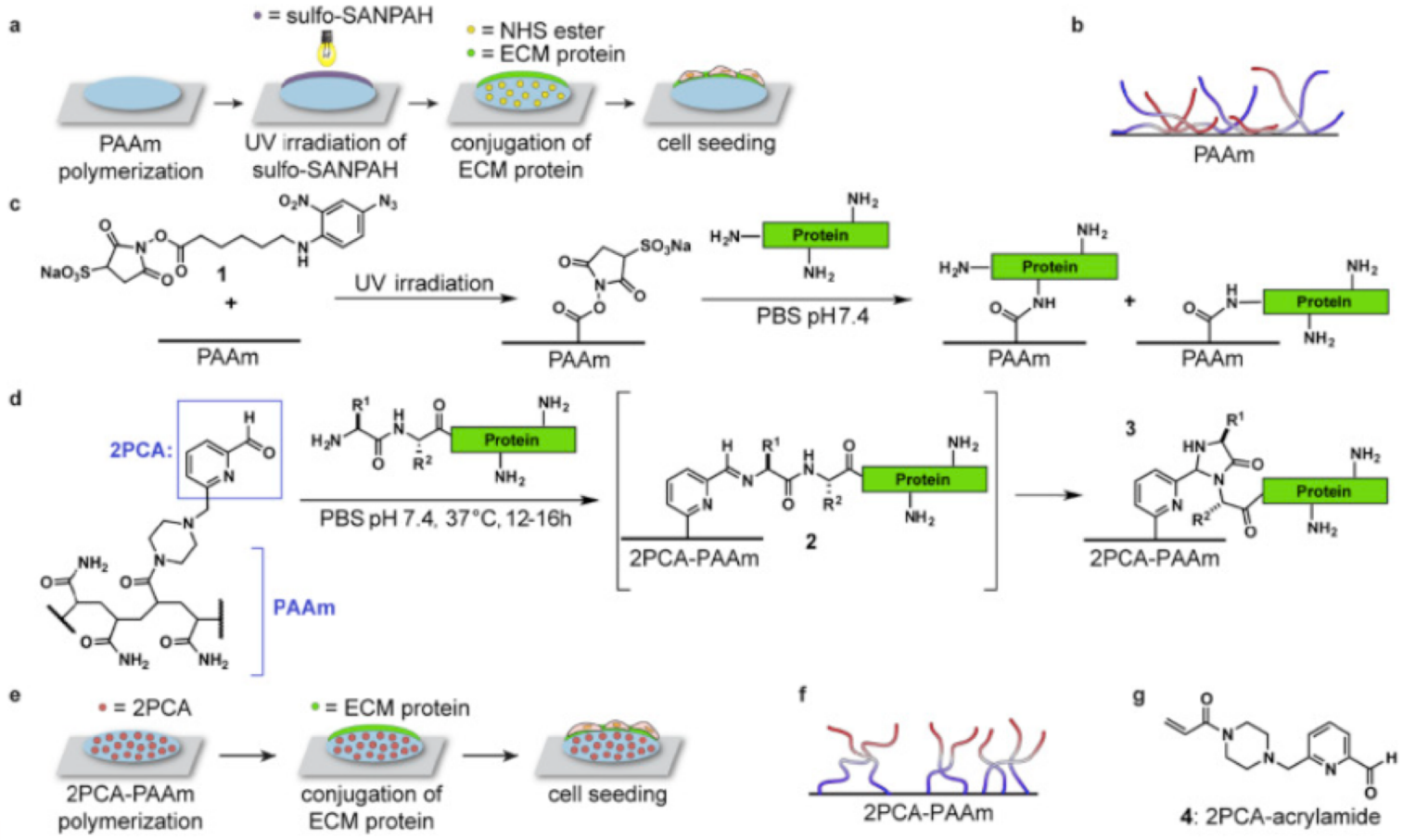
Figure 1. General scheme for the selective capture of LA-modified enzymes using βCD-functionalized sepharose resin. The noncovale immobilized enzyme is released from the resin by competitive bin with βCD.

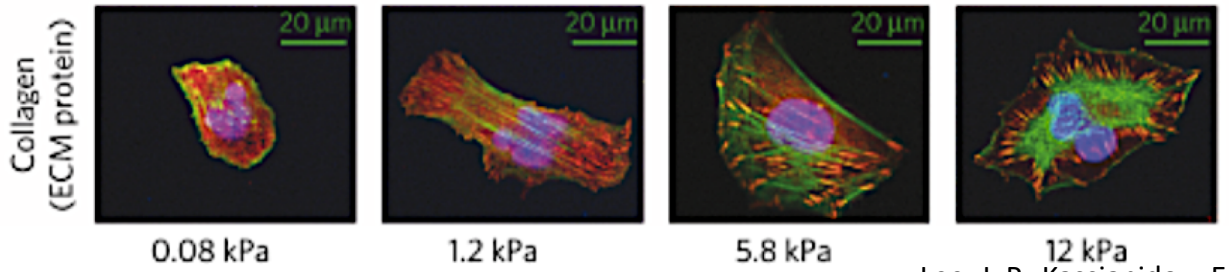
- ✓ Lithocholic acid was attached to SrtA by 2PCA-based chemistry.
- ✓ LA-SrtA is selectively captured with βCD functionalized resin.



Rosen, C. B., Kwant, R. L., MacDonald, J. I., Rao, M., & Francis, M. B. (2016). *ACIE, 55*(30), 8585-8589.

N-terminal specific conjugation of ECM (application)





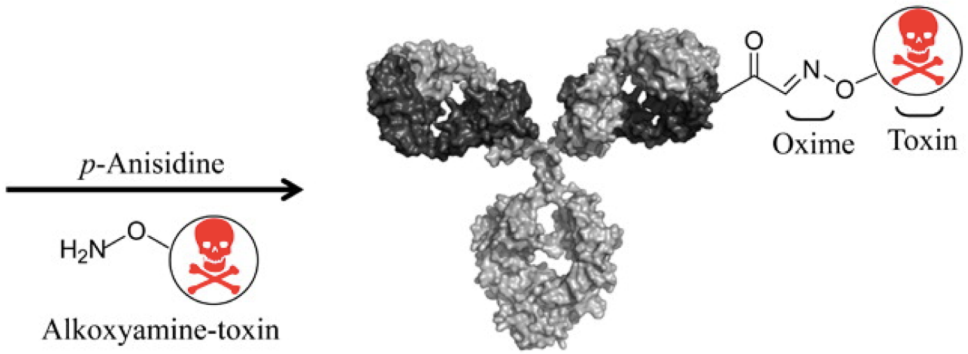
Vinculin
F-actin
DAPI

Lee, J. P., Kassianidou, E., MacDonald, J. I., Francis, M. B., & Kumar, S. (2016)., *Biomaterials, 102*, 268-276.

N-terminal modification for ADC

N-terminal serine engineered antibody

Antibody with *N*-terminal aldehyde



Antibody-drug conjugate

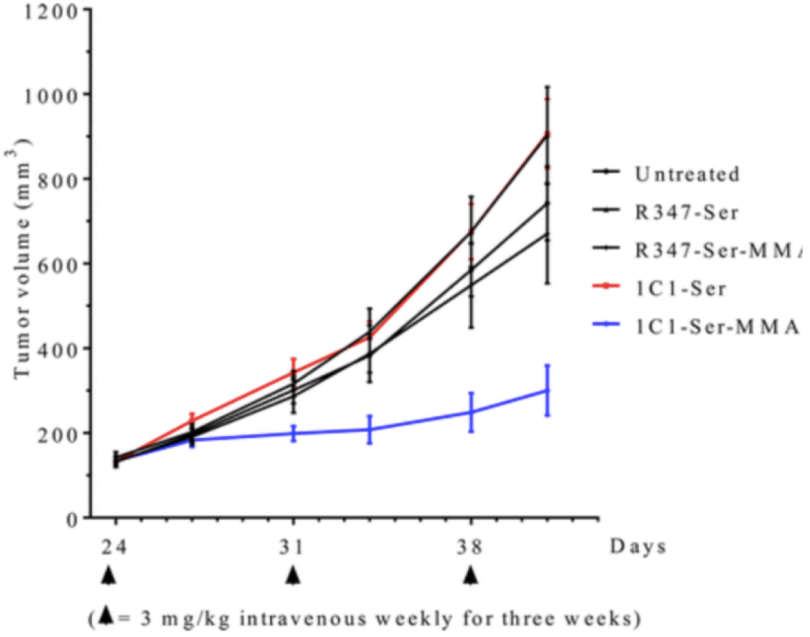


Figure 10. 1C1-Ser-MMAE has potent and specific antitumor activi in a mouse xenograft model of human prostate cancer.

Thompson, P., Bezabeh, B., Fleming, R., Pruitt, M., Mao, S., Strout, P., . . . Dimasi, N. (2015). *Bioconjugate Chemistry, 26*(10), 2085-2096.