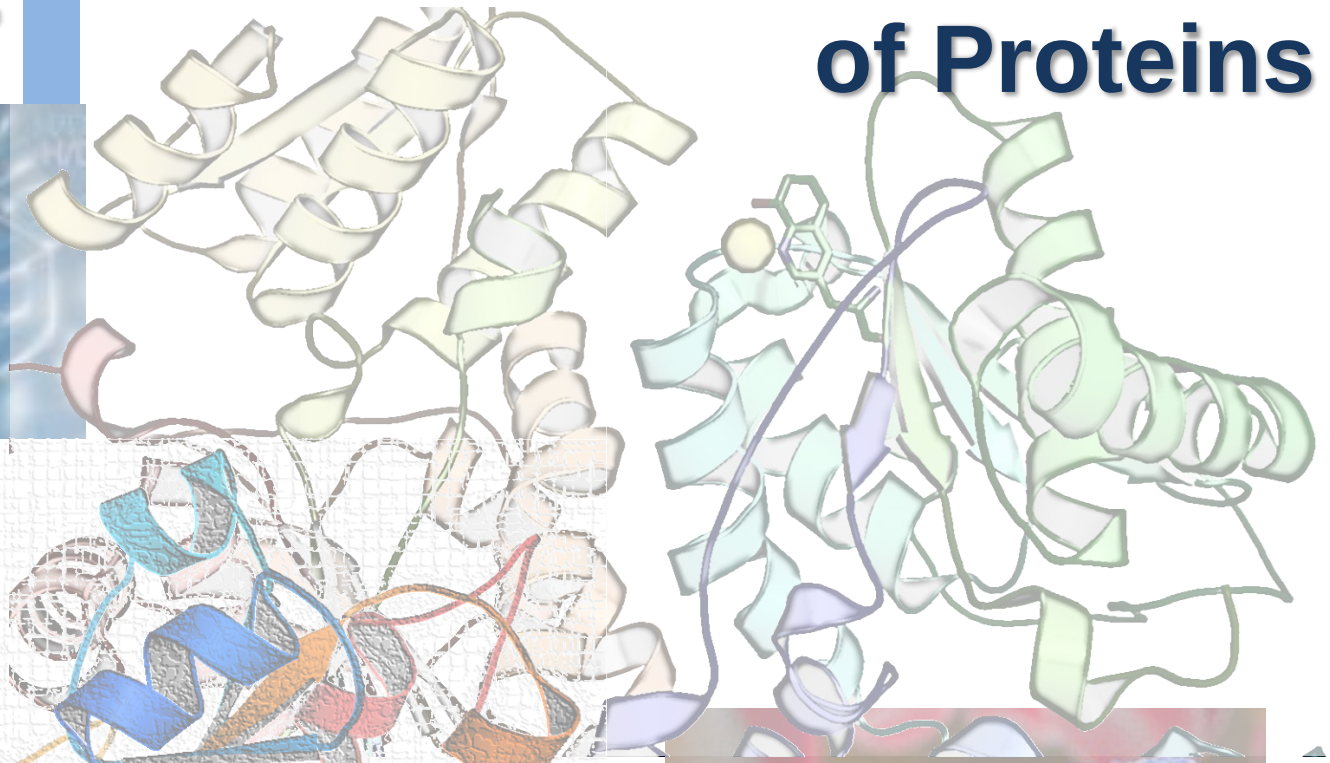
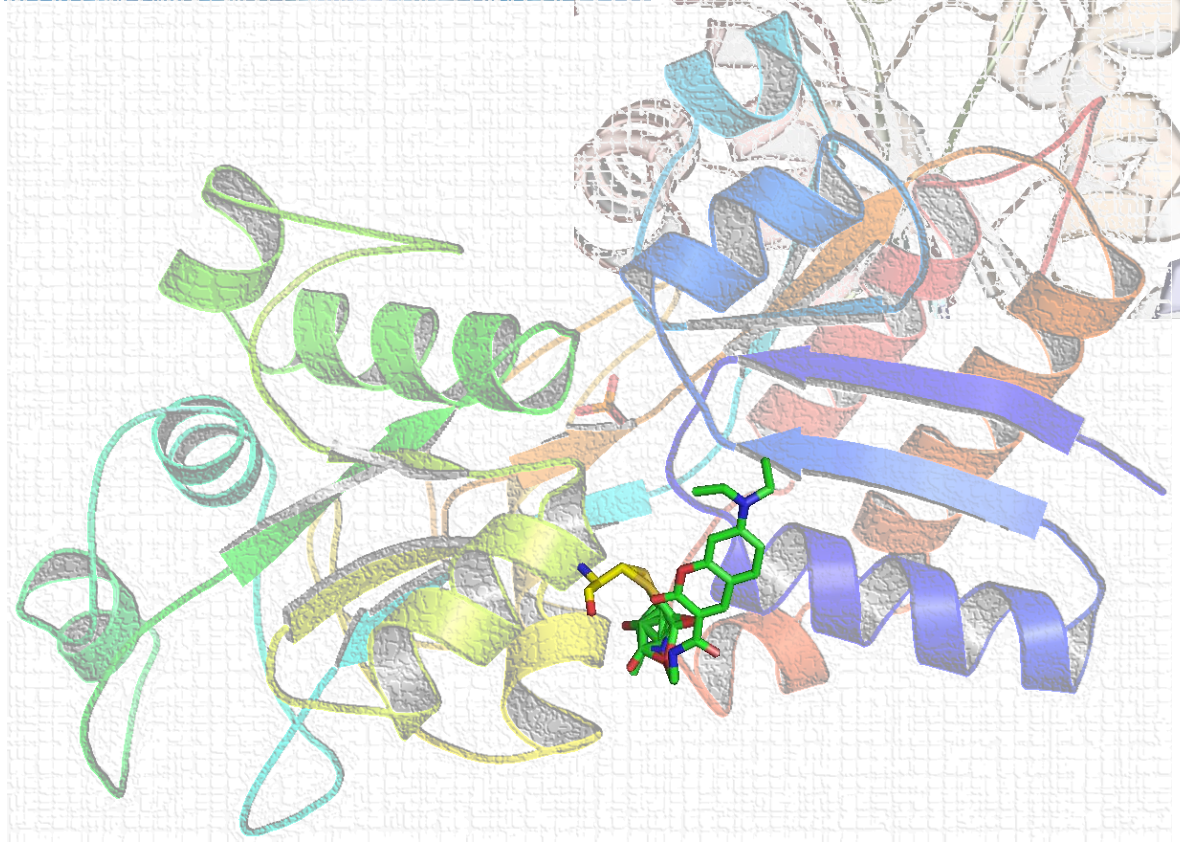
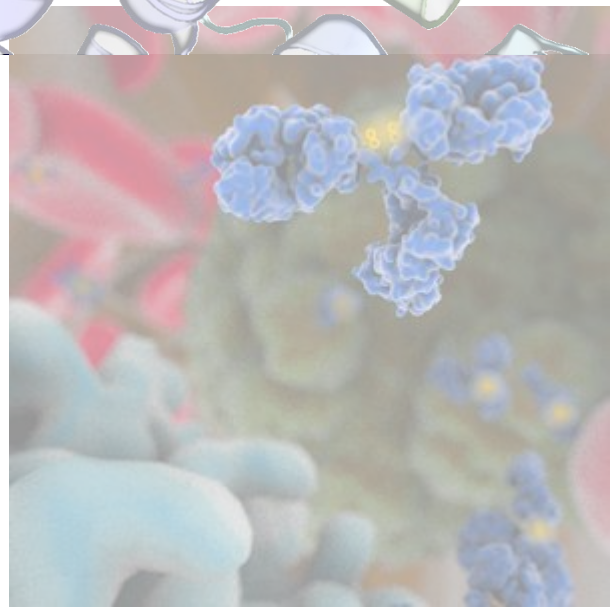
CHAPTER 4 Chemical Modificaiton of Proteins







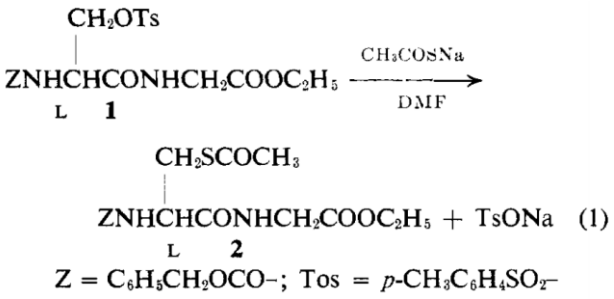
◆ Chemical protein modification

- Nature modifies proteins to increase the diversity of protein structure and function (Chapter 3)
- Our ability to install such modifications is essentially limited by the chemistry that is available
  - Reaction at a single amino acid or site among a sea of reactive groups is a significant and exciting challenge in both chemo- and regioselectivity
  - Potential transformations are molded by the need for biologically ambient conditions (that is, < 37 ℃, pH 6–8, aqueous solvent)
  - The modification should proceed with near total conversion to generate homogeneous constructs

◆ Classical examples

Conversion of the L-Serine Residue to an L-Cysteine Residue in Peptides*

Christine Zioudrou,† Meir Wilchek, and Abraham Patchornik

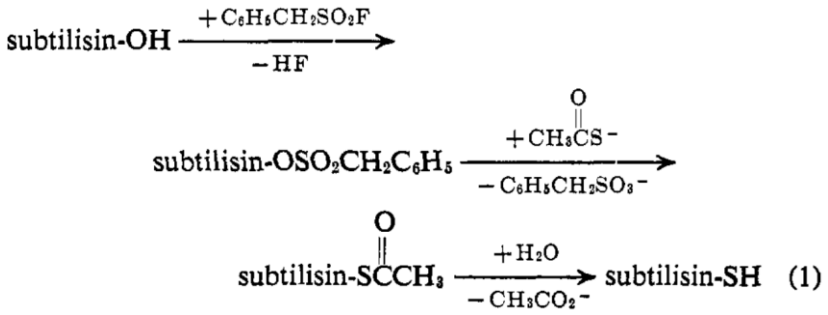


◆ This paper is pivotal in priming the concept of convergent protein synthesis; the ability to selectively ′chemically mutate′/′post-expressionally mutating′ one residue to another has had long standing implications.

Biochemistry **1965**, *4*, 1811.

A New Enzyme Containing a Synthetically Formed Active Site. Thiol-Subtilisin[1]

Laszlo Polgar, Myron L. Bender
Division of Biochemistry of the Department of Chemistry
Northwestern University, Evanston, Illinois 60201
Received March 11, 1966



◆ An enzyme (protein) was mutated by a chemical method for the first time.

JACS **1966**, *488*, 3153.

◆ Classical examples

*THE CONVERSION OF SERINE AT THE ACTIVE SITE OF SUBTILISIN TO CYSTEINE: A "CHEMICAL MUTATION"**

By Kenneth E. Neet and D. E. Koshland, Jr.

DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated September 26, 1966

TABLE 1

COMPARISON OF THE ACTIVITY OF SUBTILISIN AND THIOL-SUBTILISIN

Substrate	Initial rate (ΔA/min/μg) Subtilisin	Initial rate (ΔA/min/μg) Thiol-subtilisin	Relative rate (%) 100× thiol-subtilisin/ subtilisin
NPA[a]	9.0×10^{-4}	3.0×10^{-4}	33
ATEE[b]	2.8×10^{-3}	$<6 \times 10^{-6}$ (footnote *c*)	<0.2
TAME[d]	7.0×10^{-5}	$<1 \times 10^{-6}$	<1.4
GPNA[e]	1.4×10^{-6}	$<7 \times 10^{-9}$ (footnote *f*)	<0.5
Ovalbumin[g]	2.6×10^{-1}	~0[f, h]	0
Casein[g]	4.0×10^{-1}	~0[f, h]	0

PNAS **1966**, *56*, 1606.

◆ Cys alkylation

- The thiol of cysteine offers a unique reactive handle within proteins
  - The most robustly nucleophilic among the 20 canonical amino acids
  - The property is exploited extensively in nature
- Selective reaction at cysteine can be achieved (over other nucleophilic residues such as Lys and His)
- The low abundance (< 2%) of cysteine in proteins often allows for facile modification at a single site
  - Introduction or removal of Cys is possible by site-directed mutagenesis

◆ Cys alkylation

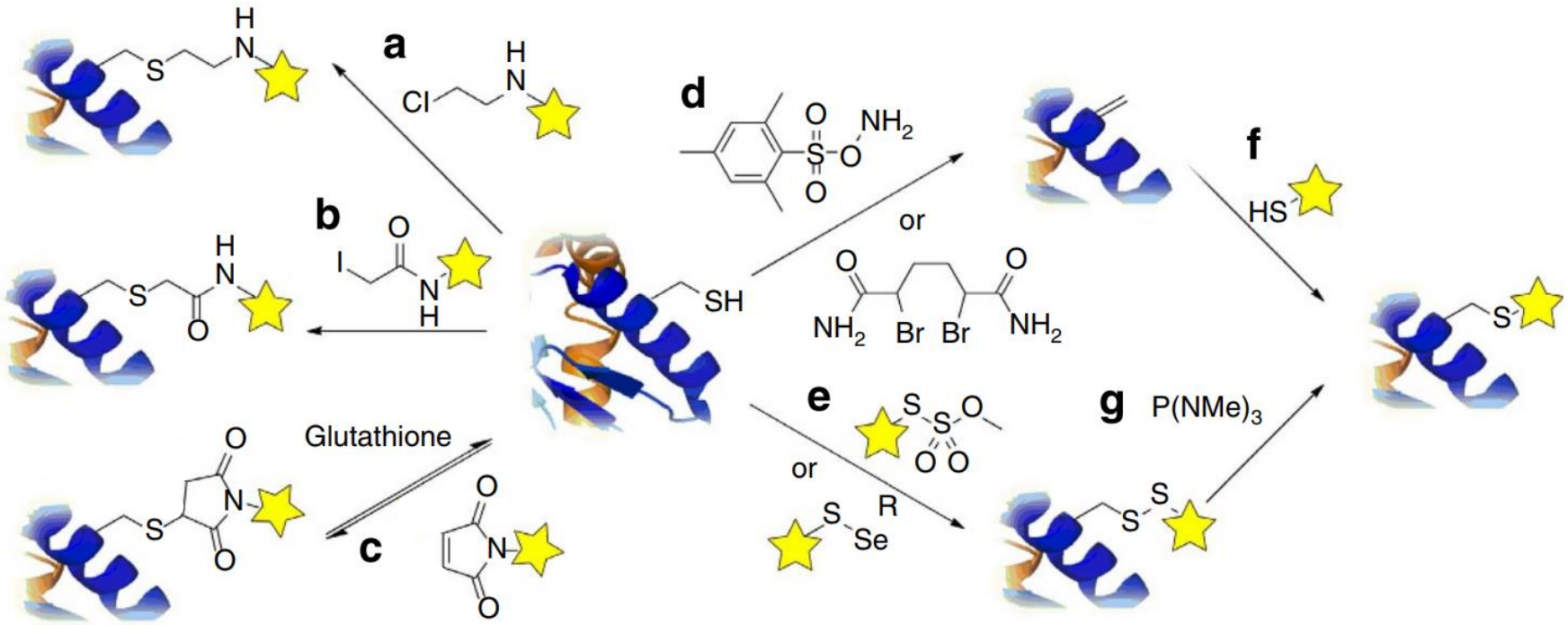


Figure 4.1 Chemical modifications at cysteine. (a) Aminoethylation (b) iodoacetamides (c) maleimides (d) Dha formation (e) disulfide formation (f) reaction of Dha with thiols and (g) desulfurization of disulfides.

Nat. Commun. **2014**, *5*, 4740.
Chem. Sci., **2011**, *2*, 1666.

◆ Cys alkylation

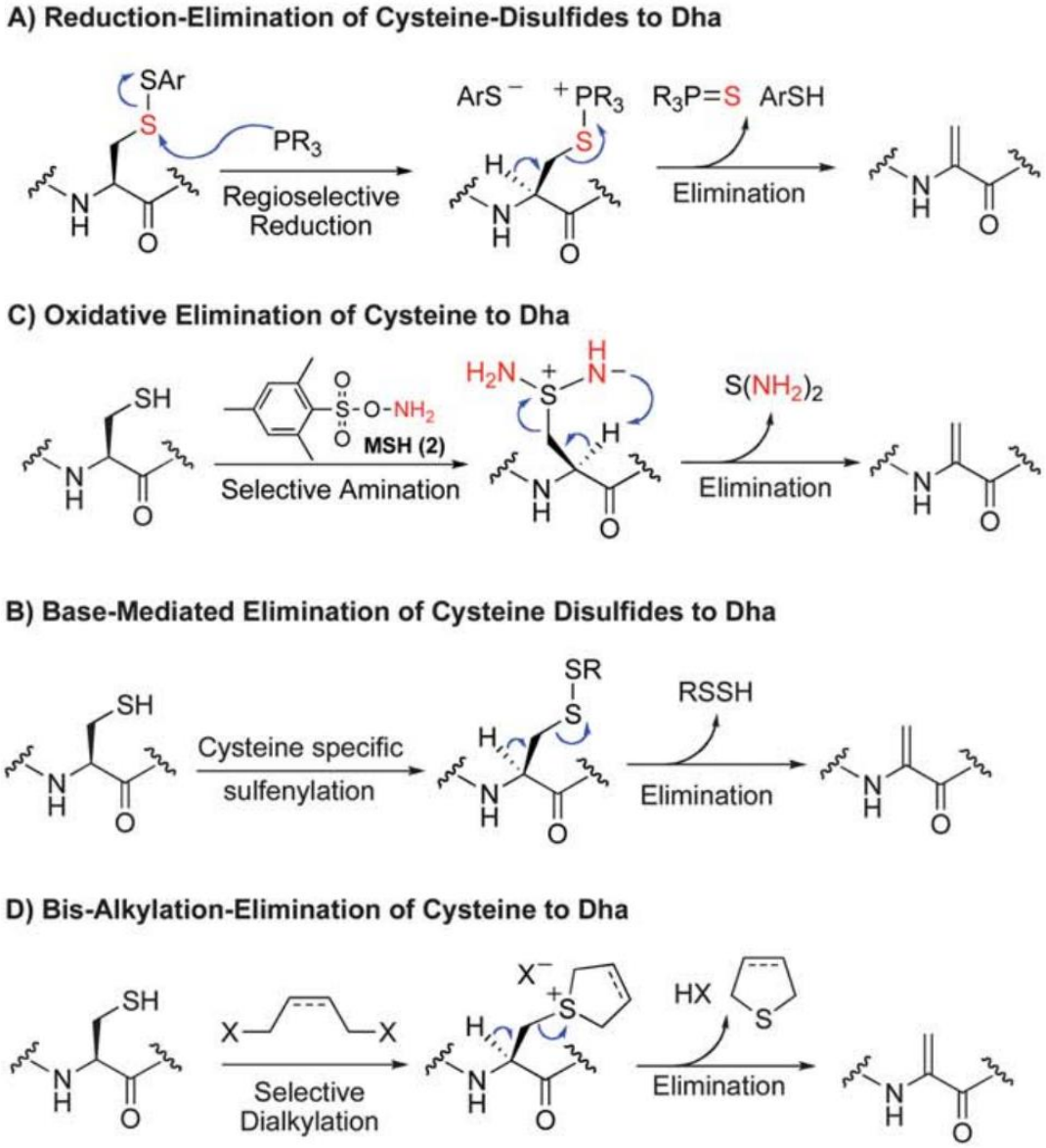


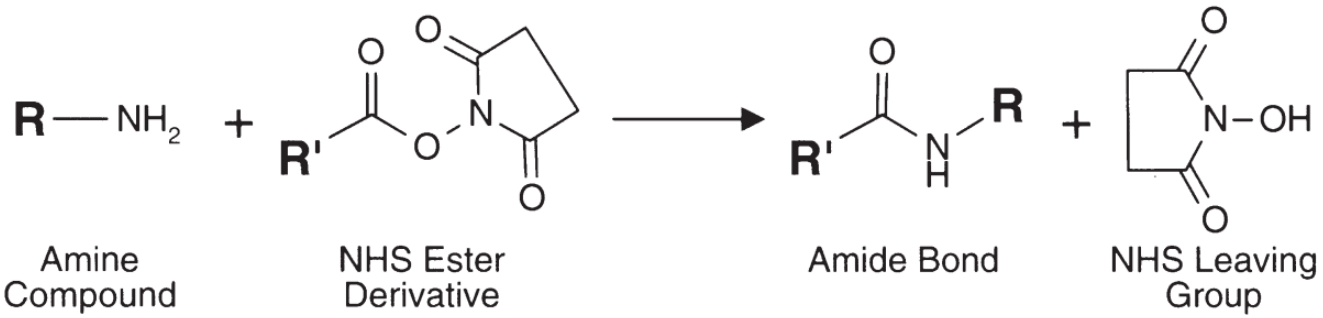
Figure 4.2
Four complementary modes of elimination of cysteine to dehydroalanine

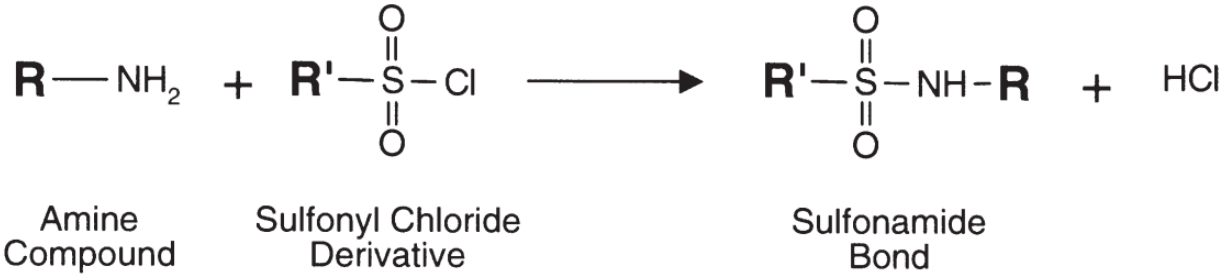
Chem. Sci. **2011**, *2*, 1666.

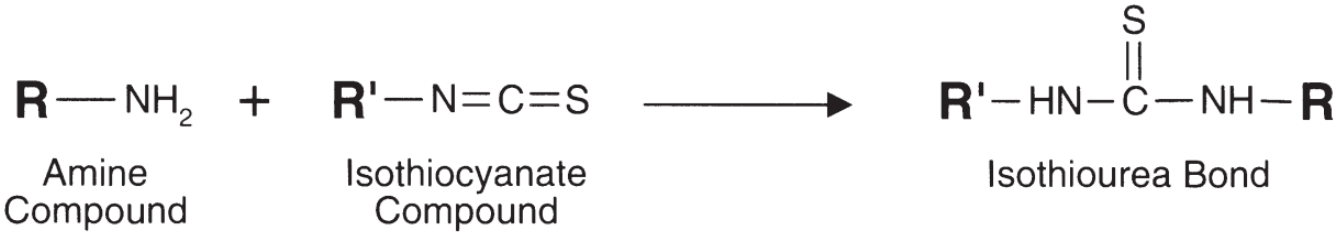
Lys conjugation

- A popular choice for protein modification along with Cys alkylation
- Selectivity can be an issue due to high natural abundance of Lys
- Preferential conjugation with amines, over the nucleophilic Cys, can be achieved through use of ‘harder’ electrophiles
  - Activated esters, sulfonyl chlorides or isothiocyanates (Scheme in the next page)
- New chemistry is being developed to improve selectivity and efficiency (*Angew. Chem. Int. Ed.* **2008**, *47*, 102.)

Lys conjugation

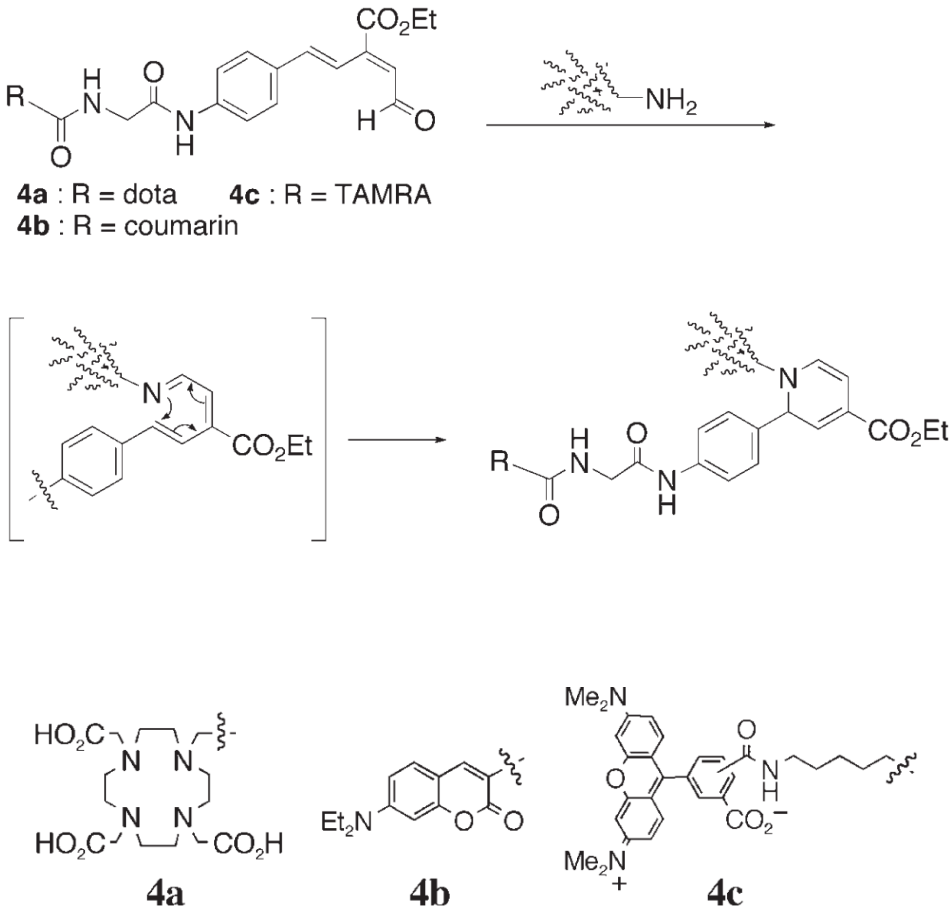






◆ Lys conjugation

- Rapid reaction with lysine through 6π-azaelectrocyclization



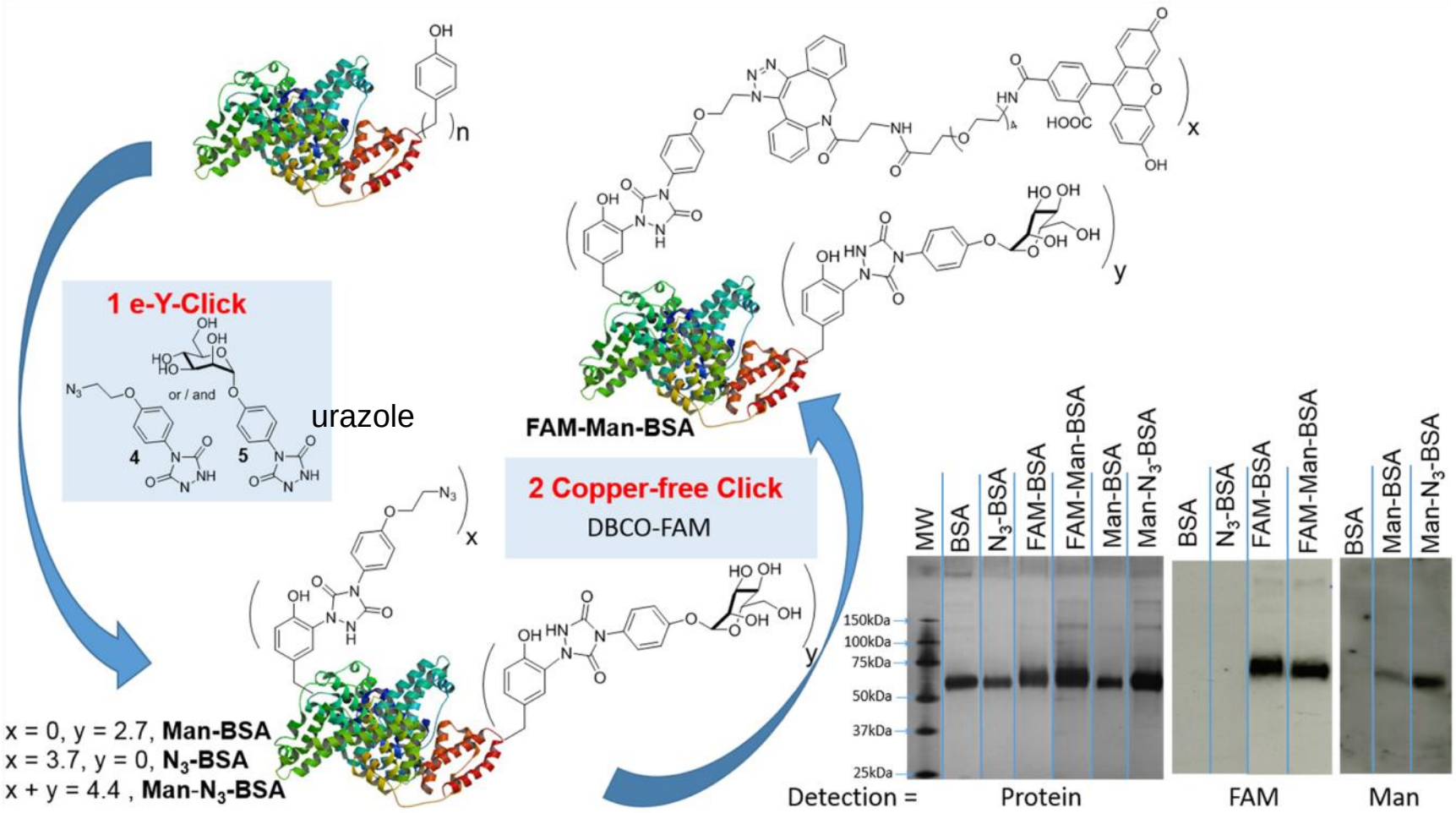
Entry	Probe	Biomolecule (*m* [μg])	Biomolecule conc [M]	Equiv probe	*t* [min]	Labeled Lys[c]
1[b]	**4a**	somatostatin (170)	5.5×10^{-4}	100	30	1
2	**4a**	albumin (100)	5.0×10^{-5}	25	10	5
3	**4a**	albumin (100)	5.0×10^{-5}	10	10	2
4	**4a**	albumin (100)	5.0×10^{-5}	500	10	20
5	**4a**	orosomucoid (62)	4.5×10^{-6}	10	30	2
6	**4a**	asialoorosomucoid (62)	4.5×10^{-6}	10	30	3
7[b]	**4a**(Gd)[d]	somatostatin (32)	3.6×10^{-4}	31	30	1
8	**4b**	albumin (120)	1.7×10^{-5}	7	30	1[e]
9	**4c**	albumin (12)	2.1×10^{-5}	7	30	1[e]
10	**4c**	anti-GFP mAb (2.0)	1.1×10^{-6}	20	30	2[e]

[a] Unless otherwise noted, reactions were performed in 0.1m phosphate buffer (pH 7.4) at 24℃. [b] in H_2O. [c] The number of labeled lysine residues was calculated by γ counting of ^{57}Co introduced to dota [d] Gd was introduced by the reaction of 4 a with 0.1m $GdCl_3$ in H_2O. [e] Estimated by emission spectra at 470 nm (coumarin) and 555 nm (TAMRA).

Angew. Chem. Int. Ed. **2008**, *47*, 102.

Tyr conjugation

- Electrochemically Promoted Tyrosine-Click-Chemistry for Protein Labeling



J. Am. Chem. Soc. **2018**, *140*, *17120*.

◆ Applications

- PEGylation
- Biotin-Streptavidin applications
- Antibody conjugations

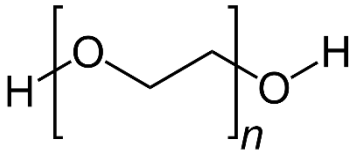
  - simple labeling, ADC, Ab-enzyme
- Fluorophore labeling

  - Fluorophore derivatives, QD
- *Chemical Ligation*

  - Original method, expressed protein ligation, intein-mediated

Polyethylene glycol (PEG)

- PEG is an organic polymer with repeating ethylene glycol units



- ***A non-immunogenic biological compound***
- PEGylation is covalent and noncovalent attachment of PEG to biological molecules such as proteins and enzymes
- PEGylation increases half-life, reduced immunogenicity, and solubility and stability
- Each PEG molecule can combine with 2-3 water molecules, making the overall compound larger and more hydrophilic

Polyethylene glycol (PEG)

- List of PEGylated compounds currently in use or approved for use

Krystexxa (PEGed uricase)	Lowers uric acid levels in order to aid in removing gout crystals.
PEGASYS (peginterferon alpha 2b)	Used with other hepatitis C medicines to treat adults with chronic hepatitis C and certain liver problems. Can be used with ribavirin to treat both adults and children with chronic hepatitis C.
Adagen (pegademase)	Modified enzyme used for Enzyme Replacement Therapy (ERT). (adenosine deaminase)
Oncaspar (pegaspargase)	Given to patients with acute lymphoblastic leukemia as part of a group of chemotherapy treatments.
Somavert (pegvisomant)	A prescription medicine for acromegaly, a disease caused by the surplus of growth hormones in the body. The goal is to have a normal IGF-1 level in the blood.
Neulasta (pegfilgrastim)	Administered to reduce the risk of infection after strong chemotherapy.
Mircera (CERA; PEG-EPO)	Used to treat symptomatic anaemia associated with chronic kidney disease (CKD).
Cimzia (certolizumab)	An injected prescription medication that works to prevent inflammation that may result from an overactive immune system.
Macugen (pegaptanib)	Utilized for the treatment of neovascular age-related macular degeneration.
Plegridy (peginterferon beta-1a)	Indicated for the treatment of patients with relapsing forms of multiple sclerosis.

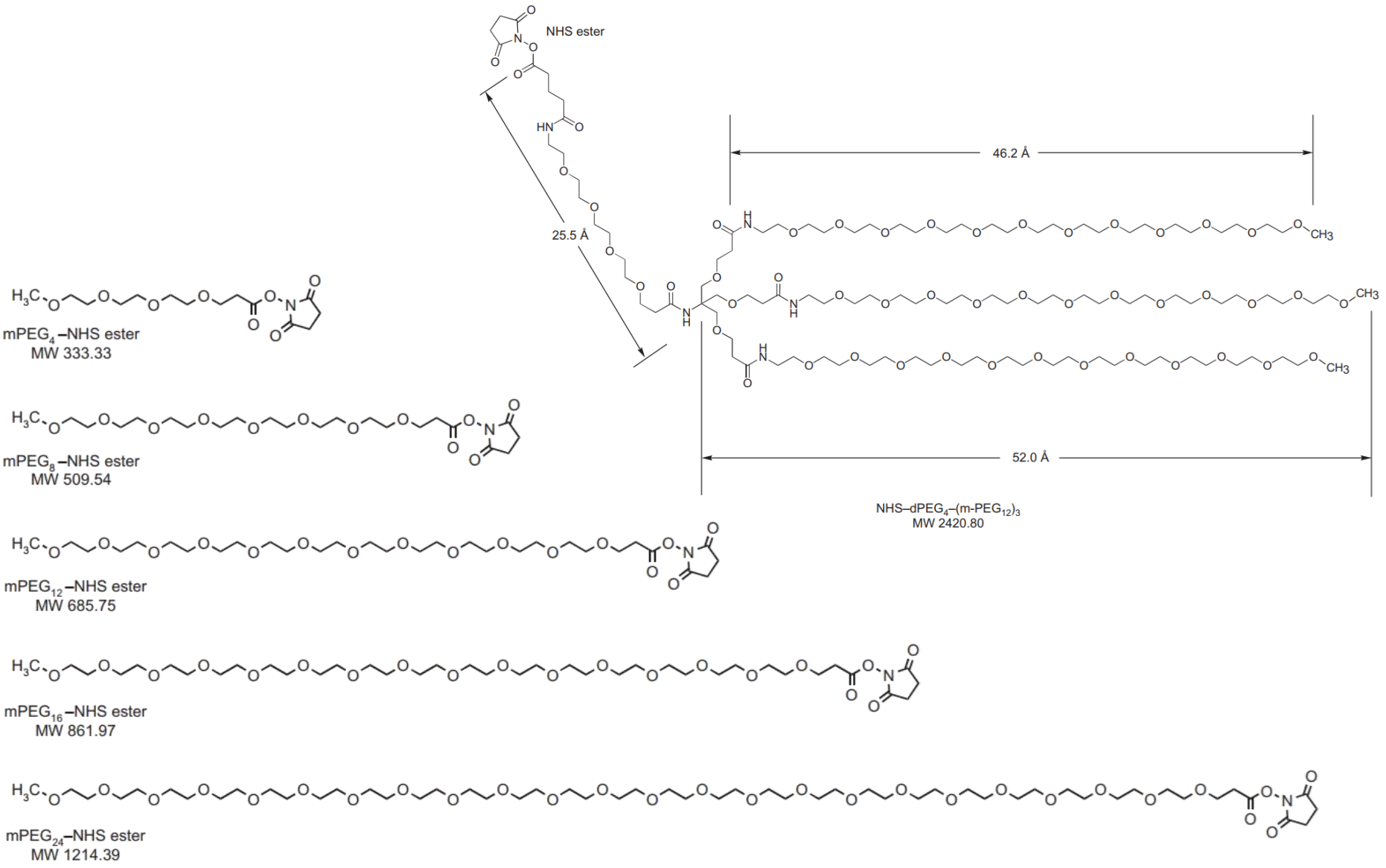
◆ Polyethylene glycol (PEG)

- List of PEGylated compounds currently in use or approved for use

Drug	Protein	Protein Size (kDa)	PEG Size (kDa)	Functional Group on PEG	Site of Attachment	Site-Specific	Year of Approval	Use
Adagen®	Adenosine deaminase	40	5	Succinimidyl ester	Lysines, serines, tyrosines, histidines	No	1990	Severe combined immunodeficiency disease
Oncaspar®	Asparaginase	31	5	Succinimidyl ester	Lysines, serines, tyrosines, histidines	No	1994	Leukemia
PegIntron®	Interferon-α-2b	19.2	12	Succinimidyl ester	Lysines, serines, tyrosines, histidines,cysteines	No, but 47.8% one isomer	2000	Hepatitis C
Pegasys®	Interferon-α-2a	19.2	40	Succinimidyl ester	Lysines	No	2001	Hepatitis C
Neulasta®	Granulocyte colony-stimulating factor (G-CSF)	18.8	20	Aldehyde	N-Terminal amine	Yes	2002	Neutropenia
Somavert®	Human growth hormone (hGH)	22	5	Succinimidyl ester	Lysines, N-terminus, phenylalanine	No	2003	Acromegaly
Mircera®	Erythropoietin	30 (18 unglycosylated)	40	Succinimidyl ester	Lysines	No	2007	Anemia
Cimzia®	Anti-tumor necrosis factor (TNF) α Fab'	51	40	Maleimide	C-Terminal cysteine	Yes	2008	Rheumatoid arthritis, Crohn disease, psoriatic arthritis
Krystexxa®	Urate oxidase	34	10	*p*-Nitrophenyl carbonate ester	Lysines	No	2010	Gout
Omontys	Synthetic, dimeric peptide (erythropoiesis stimulating agent)	4.9	40 (2 20 kDa chains)	Succinimidyl ester (added during chemical synthesis of the peptide)	Lysines	Yes	2012 (Recalled 2014)	Anemia in chronic kidney disease

Int. J. Mol. Sci. **2015**, *16*, 25831.

Reagents for PEGylation – NHS esters

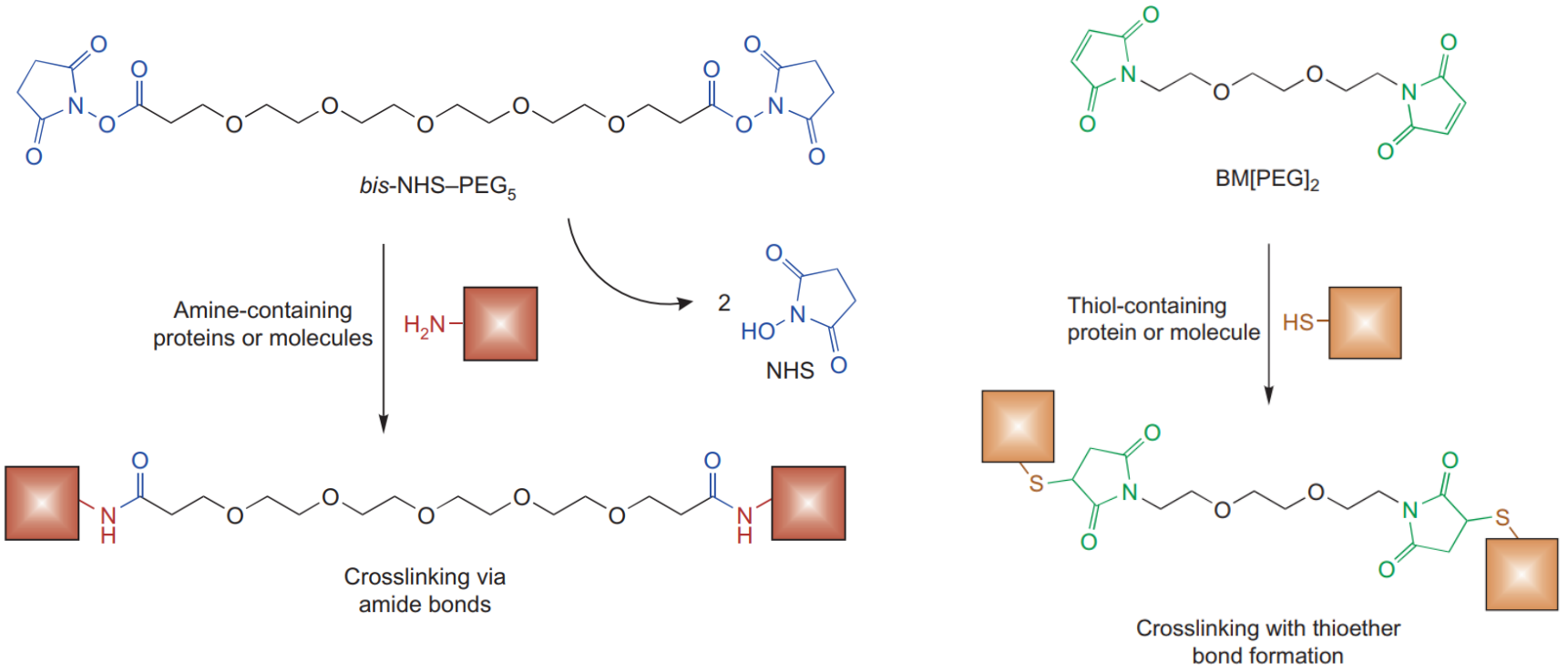


Reagents for PEGylation – NHS esters

- The reaction of the NHS esters occurs at physiological pH or under slightly basic conditions to couple rapidly with amines and form amide bonds
- The NHS ester groups are also subject to hydrolysis in aqueous solution
  - The rate of hydrolysis increases with increasing pH
  - The more hydrophilic the molecule, the greater the potential for hydrolysis
  - The half-life of hydrolysis effectively triples upon lowering the pH one unit (depends on conditions)
  - To aid in NHS ester stability maintaining a reaction pH in the range of 7.0 to 7.5 is optimal for most applications

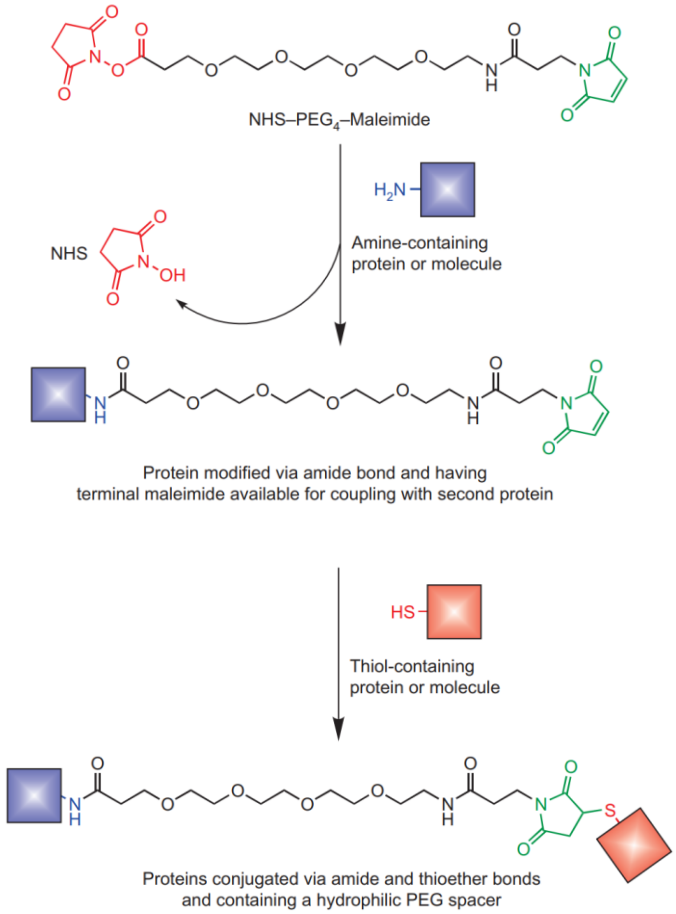
Homobifunctional PEG crosslinkers

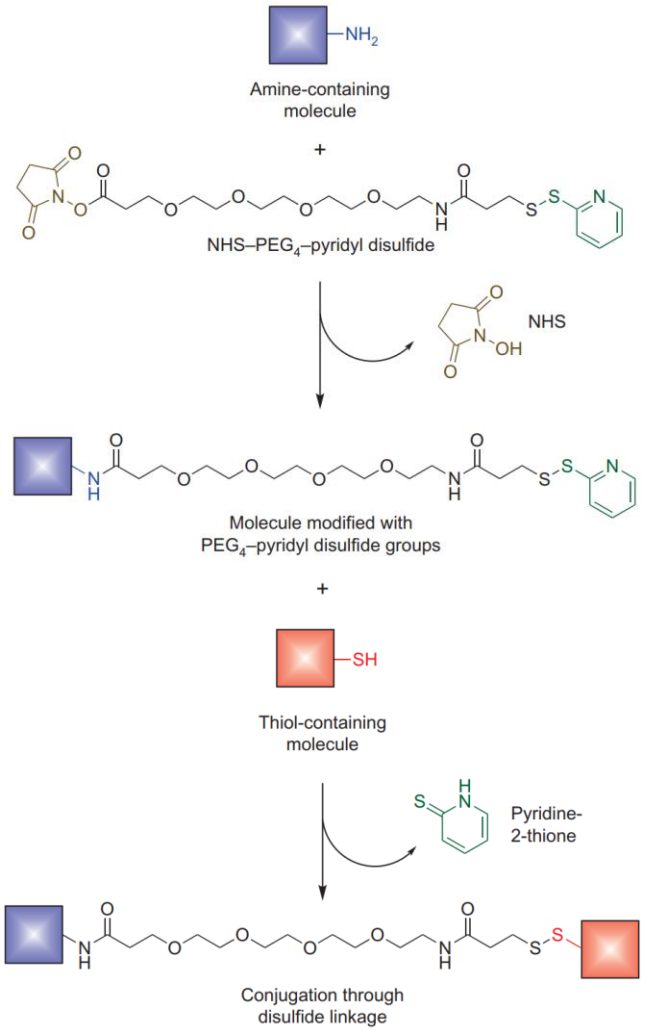
- The use of homobifunctional NHS-ester compounds allows to study protein interactions



Heterobifunctional PEG crosslinkers

- The use of heterofunctional PEG linkers can be used for protein-protein conjugation with benefits in solubility and immunogenecity





Streptavidin

- Streptavidin is a biotin-binding protein originated from *Streptomyces avidinii*
- Streptavidin's structural differences can overcome some of the nonspecific binding deficiencies of avidin
  - Not a glycoprotein, thus there is no potential for binding to carbohydrate receptors
  - pI ~ 5−6 (avidin has pI~10) reduces nonspecific binding due to ionic interaction with other biomolecules
  - These factors lead to better signal-to-noise ratios in assays using streptavidin–biotin interactions than those employing avidin–biotin

Use of streptavidin-biotin interaction in assay systems

- There are several basic immunoassay designs that make use of the enhanced sensitivity afforded by the (strept)avidin–biotin interaction

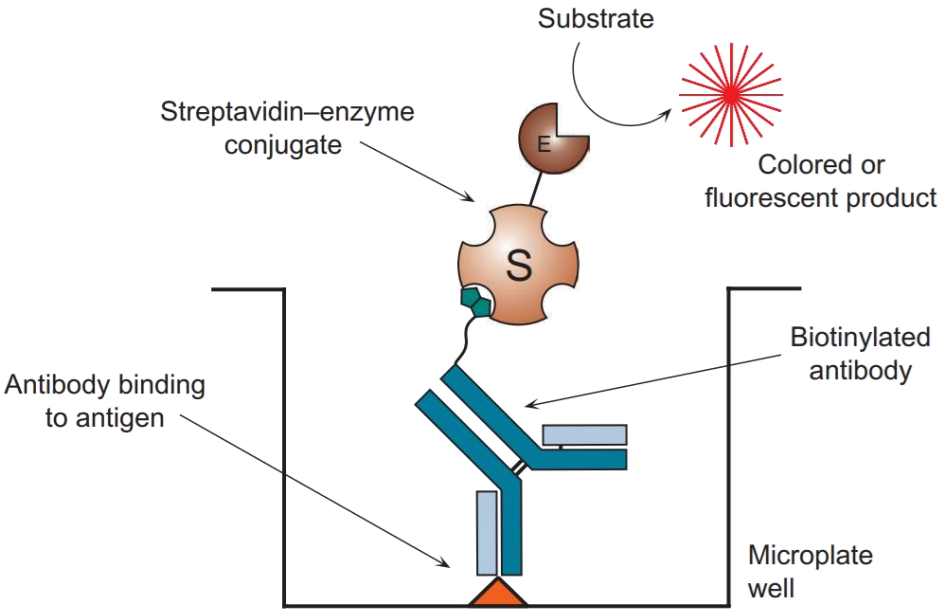


FIGURE 4.1 The basic design of the labeled avidin–biotin (LAB) assay system.

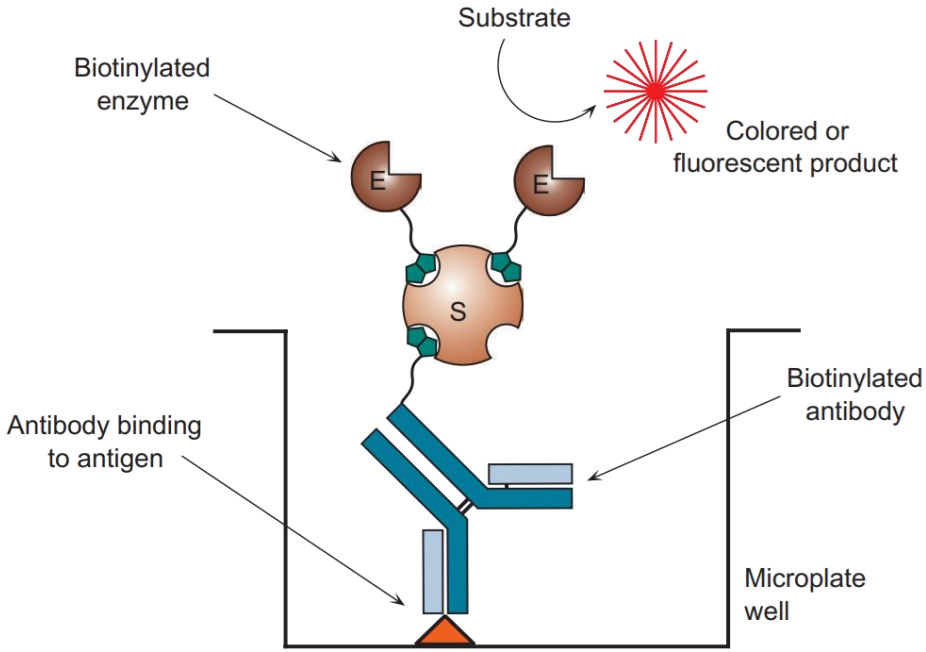


FIGURE 4.2 The basic design of the bridged avidin–biotin (BRAB) assay system.

Use of streptavidin-biotin interaction in assay systems

- The ABC system (for avidin–biotin complex) increases the detectability of antigen
- It is a common assay technique used for microplate-based ELISA assays and for immunohistochemistry (IHC) procedures

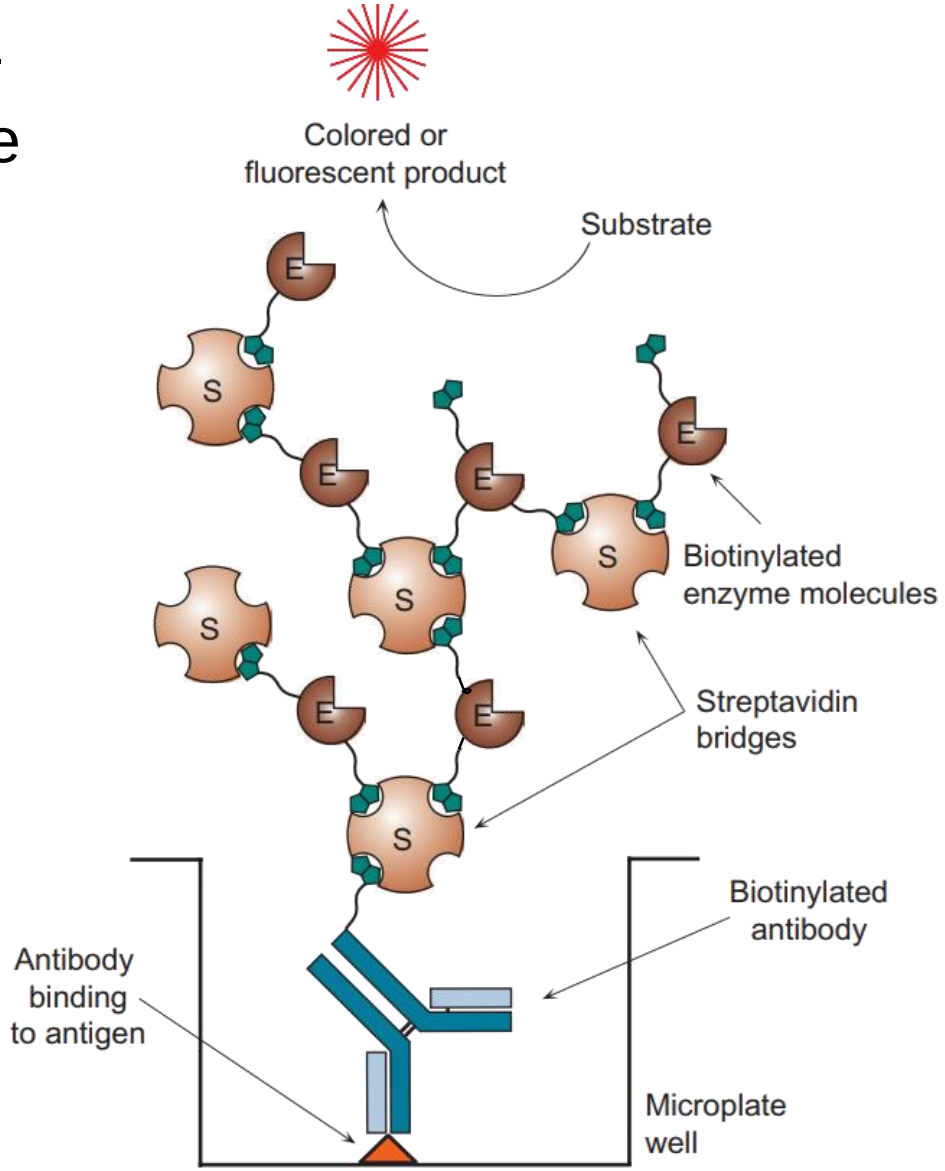


FIGURE 4.3 The assay design of the avidin–biotin complex (ABC) system.

Preparation of streptavidin conjugates

- In most applications, streptavidin is conjugated with enzymes such as HRP and AP

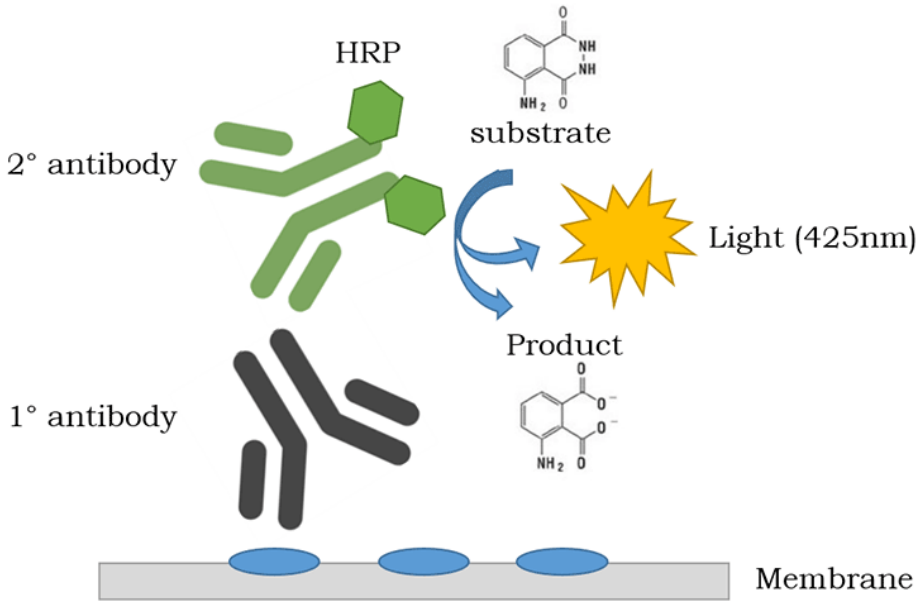


FIGURE 4.4 Chemiluminescent reaction of HRP substrate-luminol. The oxidation of luminol by peroxide results in creation of an excited state product called 3-aminophthalate. This product decays to a lower energy state by releasing photons of light.

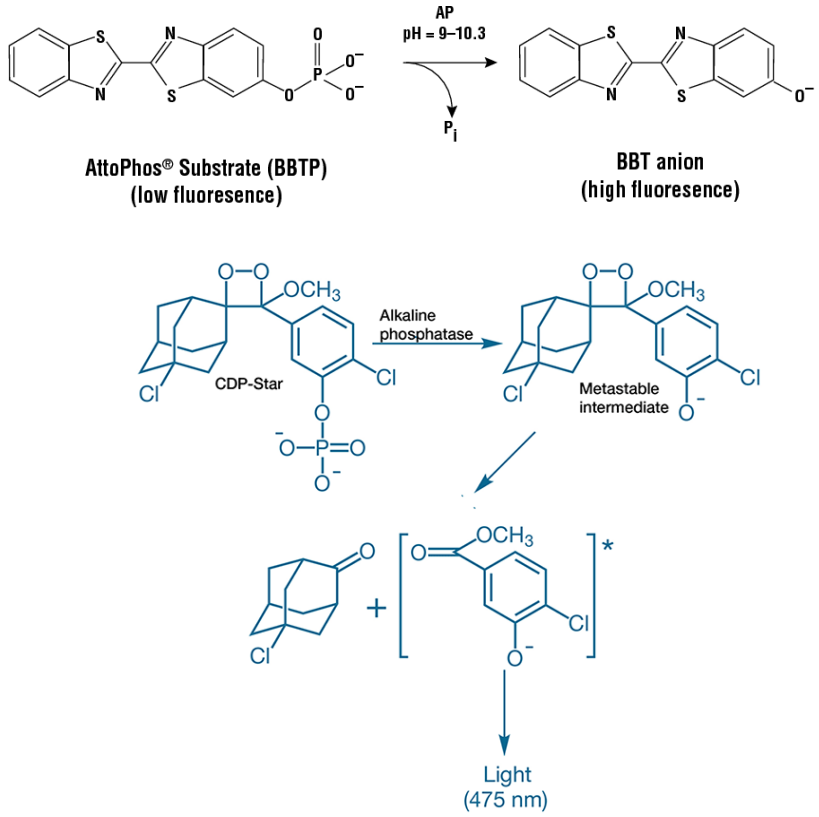


FIGURE 4.5 Schemes for visualization of AP reactions used in biochemical assays

Preparation of streptavidin conjugates

- A common protein coupling reagent (heterobifunctional)

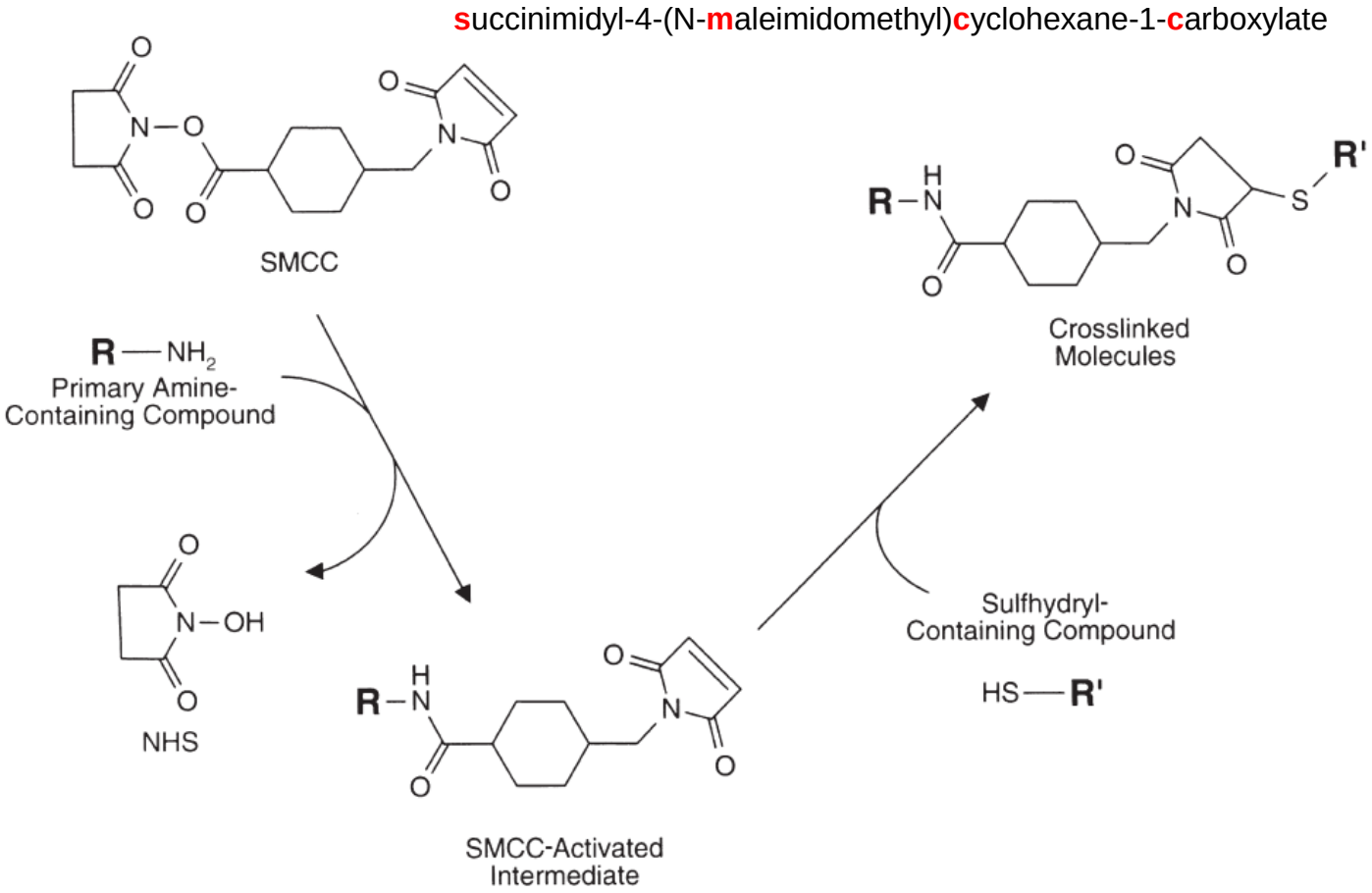


FIGURE 4.6 SMCC reacts with amine-containing molecules to form stable amide bonds. Its maleimide end may then be conjugated to a sulfhydryl-containing compound to create a thioether linkage.

Preparation of streptavidin conjugates

Streptavidin has no free Cys

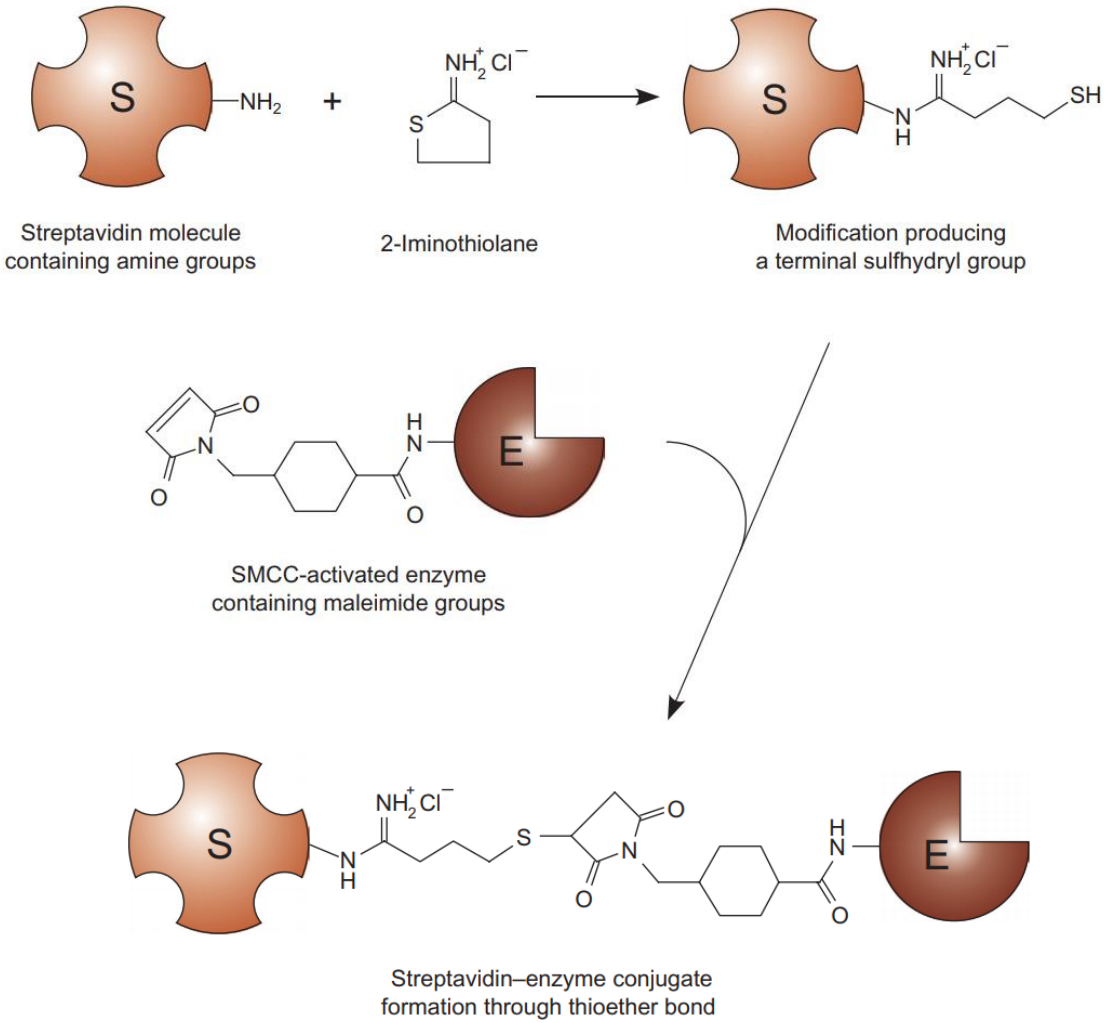


FIGURE 4.7 Streptavidin can be modified with 2-iminothiolane (Traut's reagent) to produce sulfhydryl groups. Subsequent reaction with a maleimide-activated enzyme produces a thioether-linked conjugate.

Preparation of streptavidin conjugates

HRP and avidin are glycoproteins

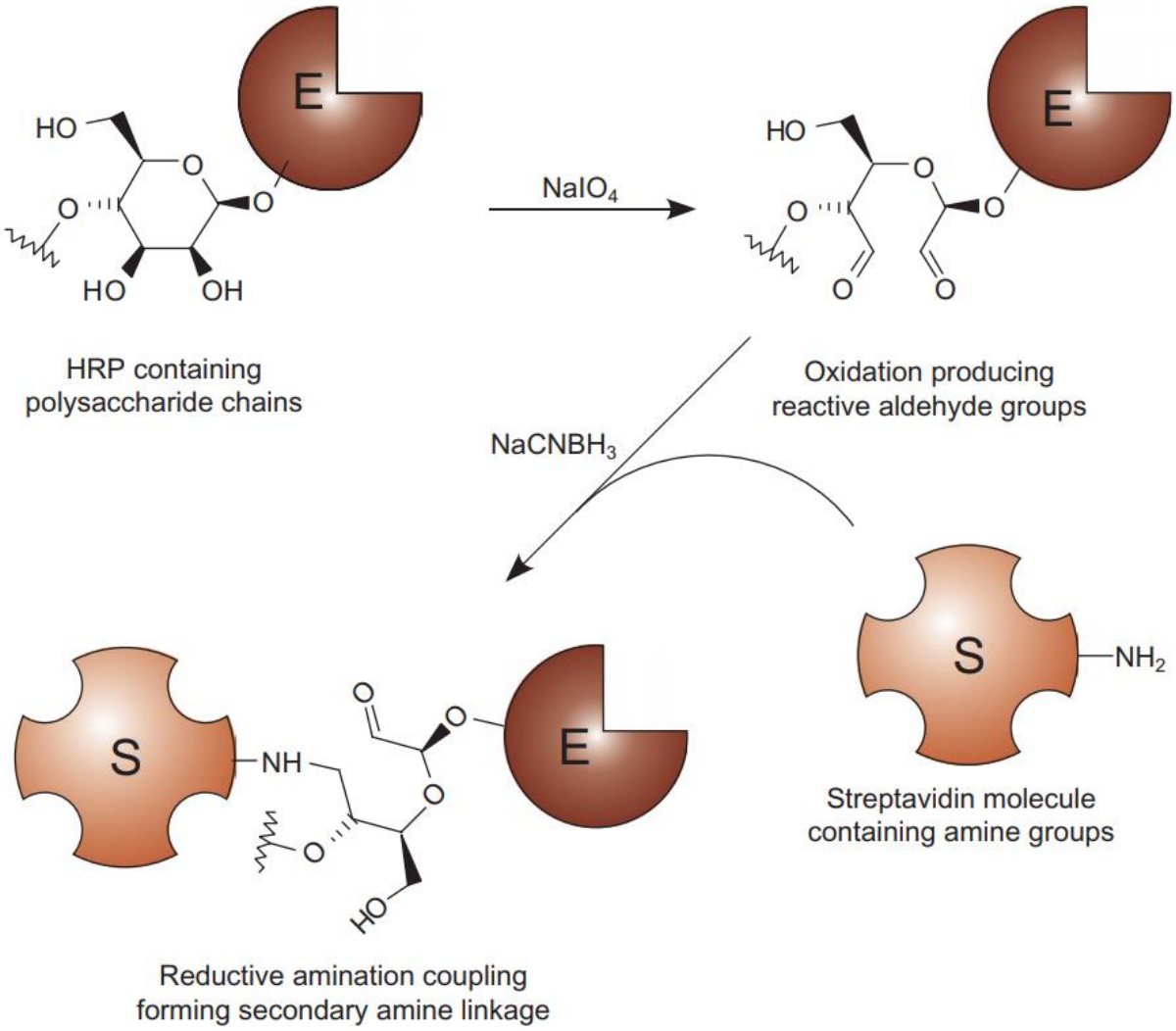


FIGURE 4.8 Oxidation of the polysaccharide components of HRP produces reactive aldehyde groups. Conjugation to streptavidin may then be achieved by reductive amination.

Fluorescence labeling of streptavidin

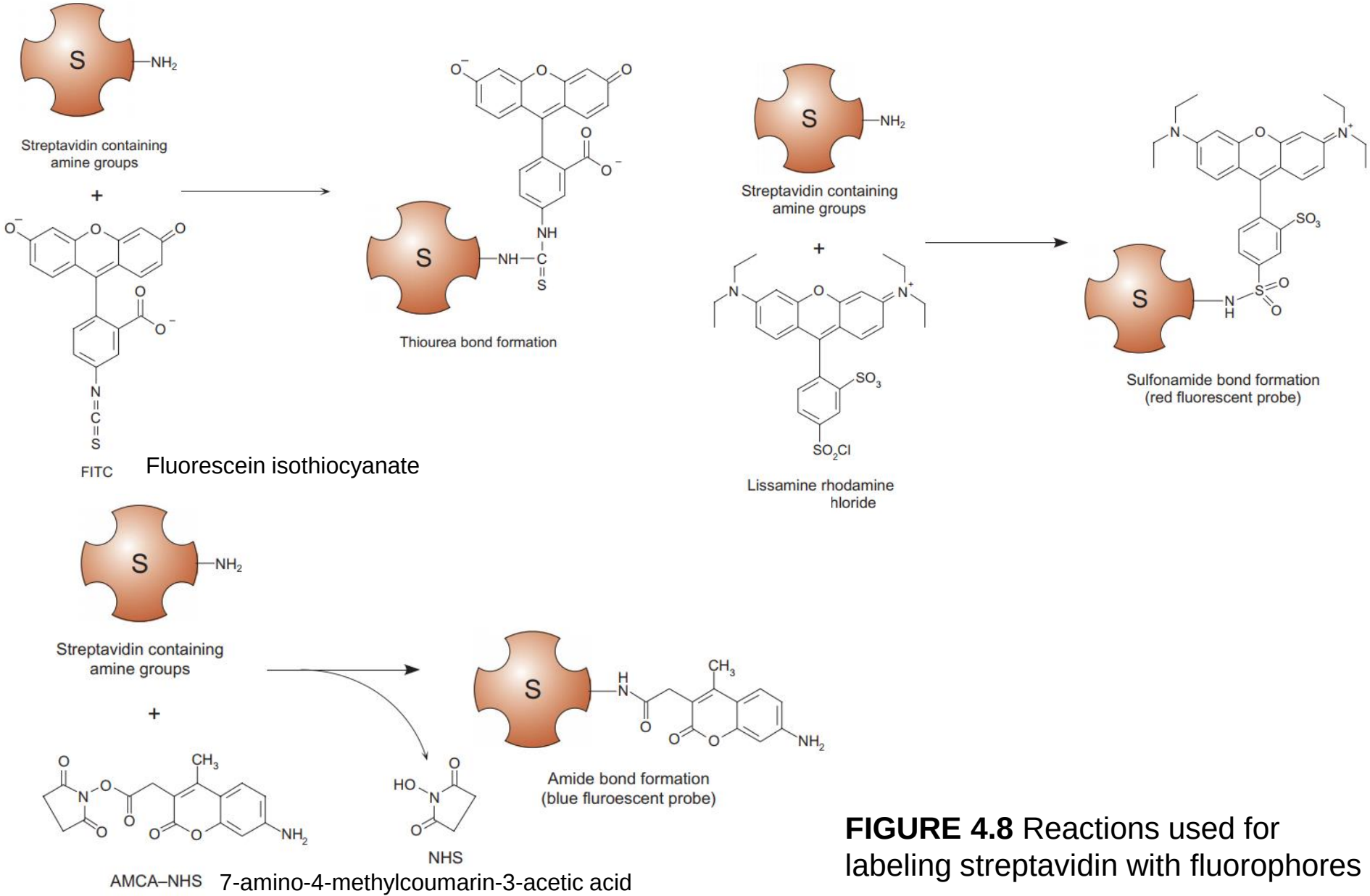
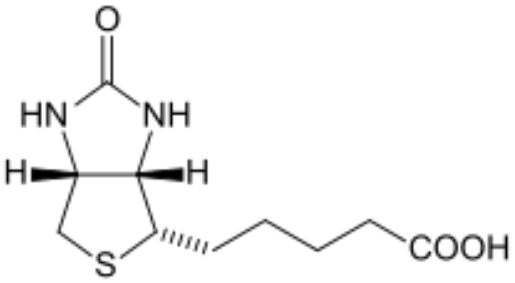


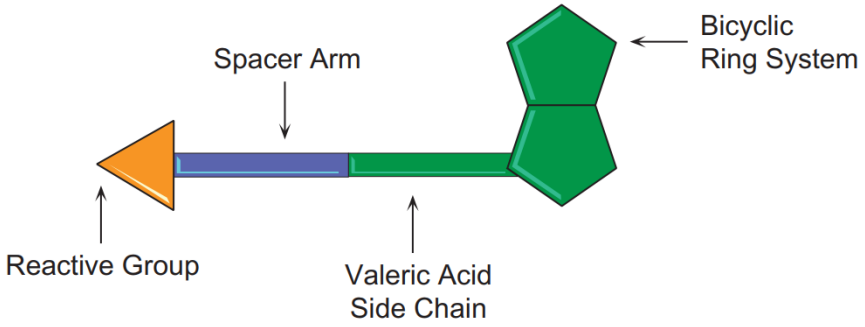
FIGURE 4.8 Reactions used for labeling streptavidin with fluorophores

Preparation of biotin conjugates

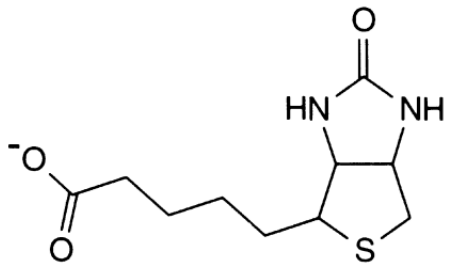
- Biotin modification of secondary molecules results in covalent derivatives containing bicyclic biotin rings extending from the parent structure.
- These biotinylation sites are still capable of binding avidin or streptavidin with the specificity and nearly the same avidity of free biotin in solution
- Since the biotin components are relatively small, macromolecules can be modified with these reagents without significantly affecting their physical or chemical properties



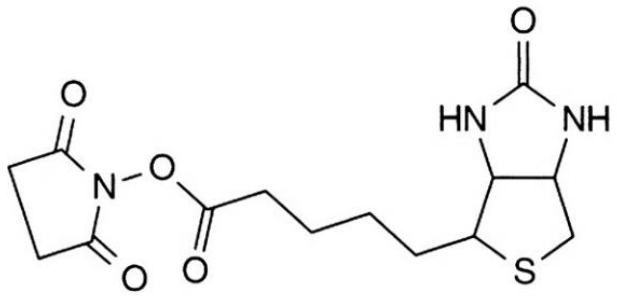
D-biotin



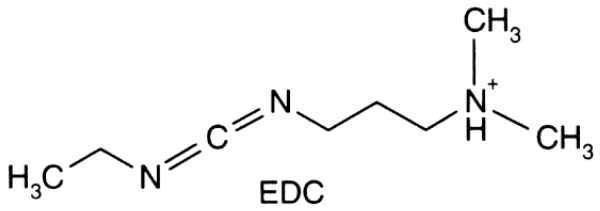
◆ Reagents for biotin conjugation



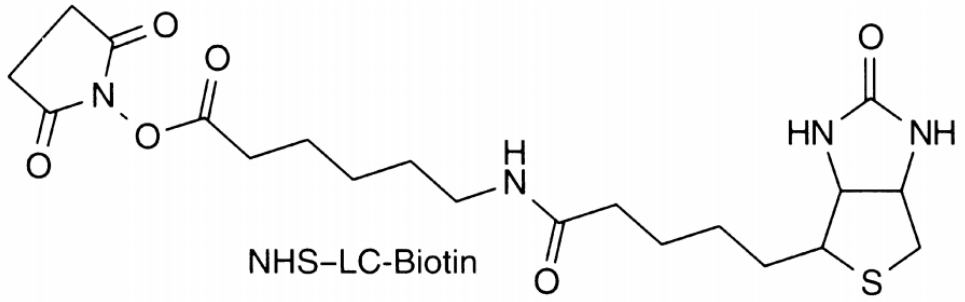
D-Biotin



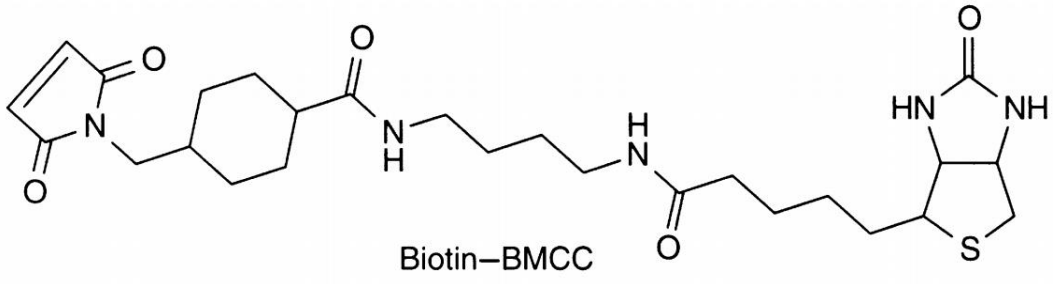
NHS–Biotin



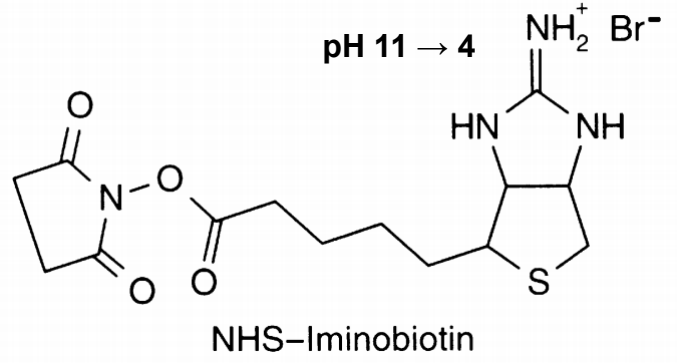
EDC



NHS–LC-Biotin



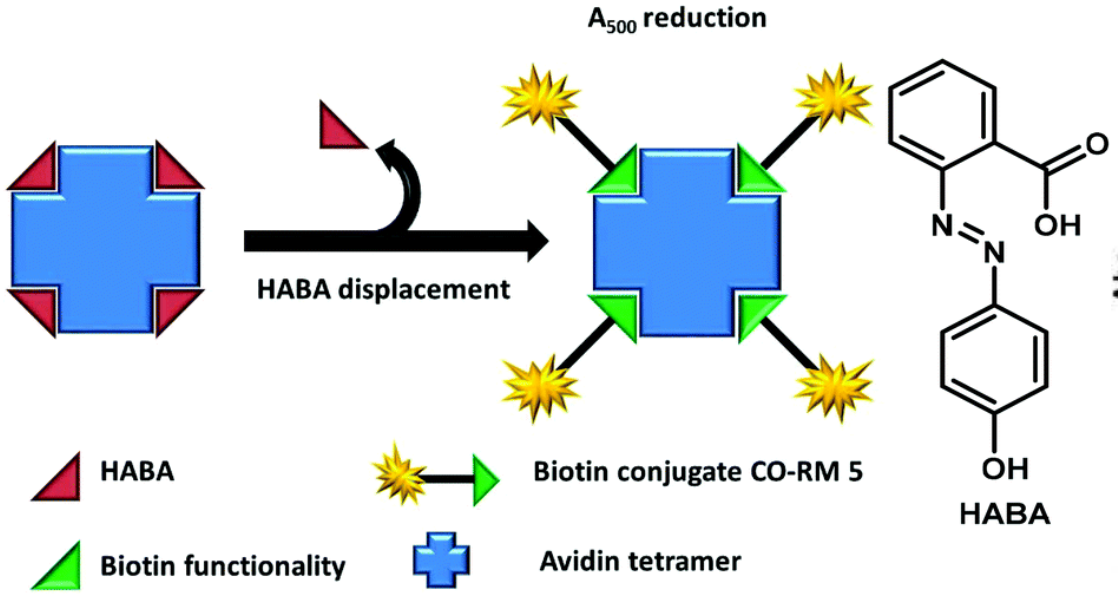
Biotin–BMCC

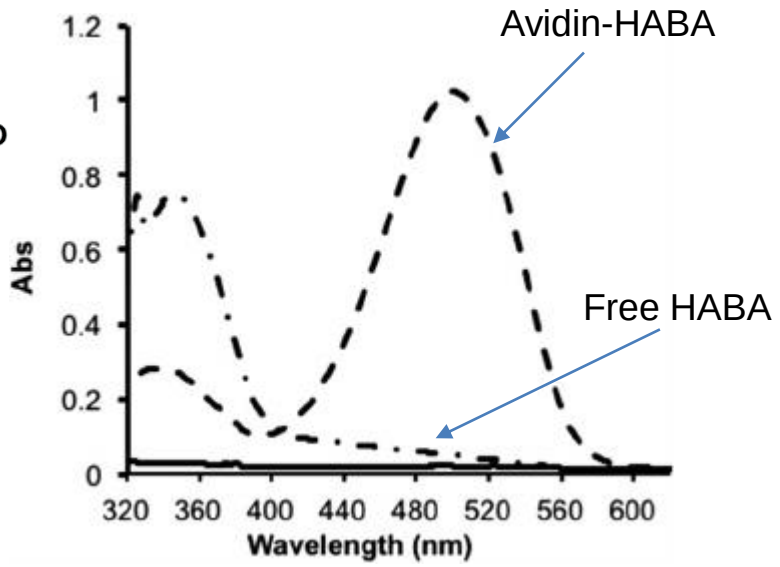


NHS–Iminobiotin

HABA dye assay

- The most common method of measuring the degree of biotinylation



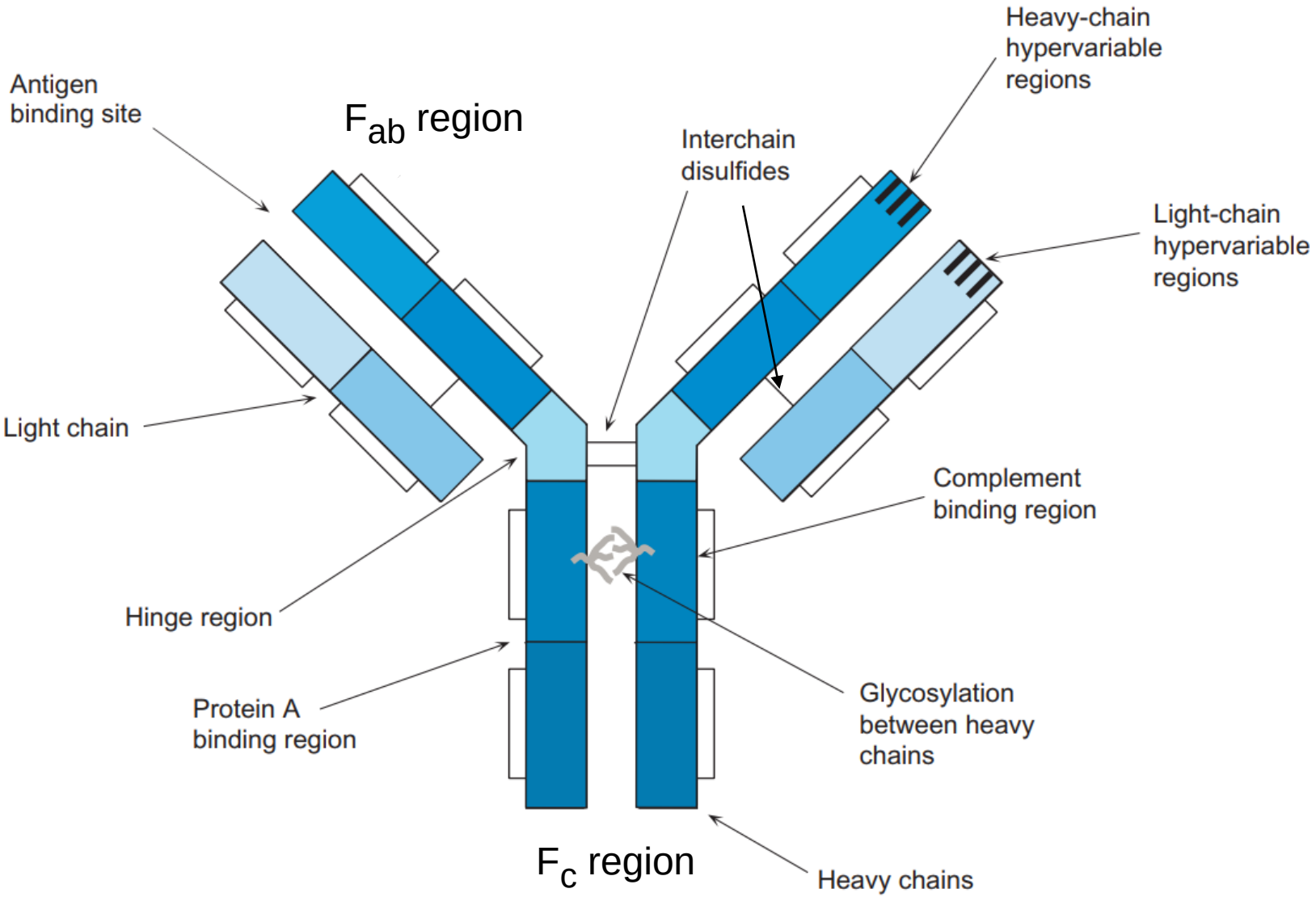


Org. Biomol. Chem. **2014**, *12*, 511.

Antibody Modificaiton

- The ability to conjugate an antibody to another protein or molecule is critically important for many applications in life science research, diagnostics, and therapeutics.
- Antibody conjugates have become one of the most important classes of biological agents associated with targeted therapy for cancer and other diseases.
- There literally are a number of markers that have been identified on tumor cells to which monoclonal antibodies have been developed for targeted therapy
- The preparation of antibody conjugates to find and destroy cancer cells in vivo has become one of the leading strategies of research into investigational new drugs

Antibody Structure

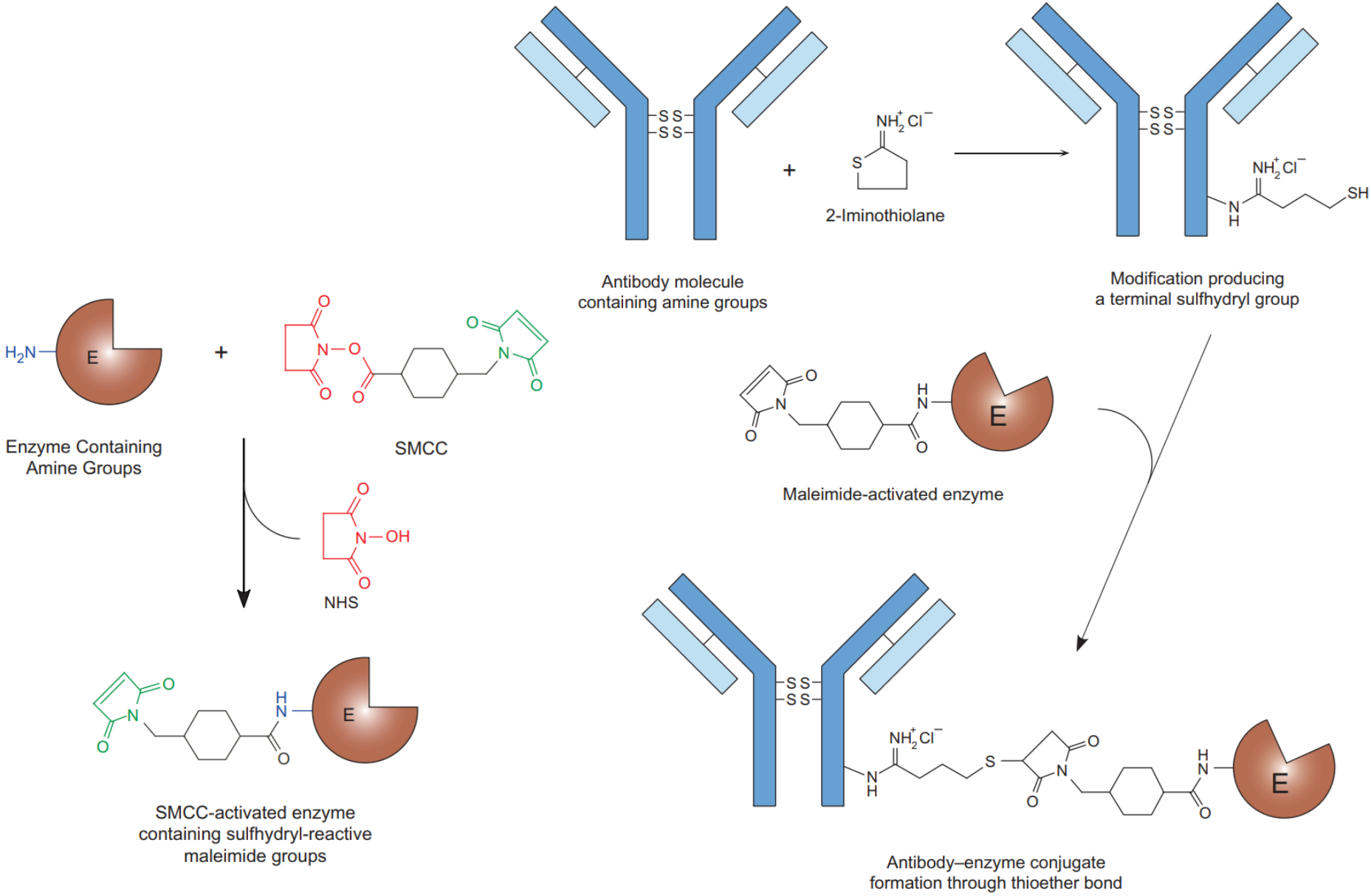


PDB 1HZH
Pymol Demo

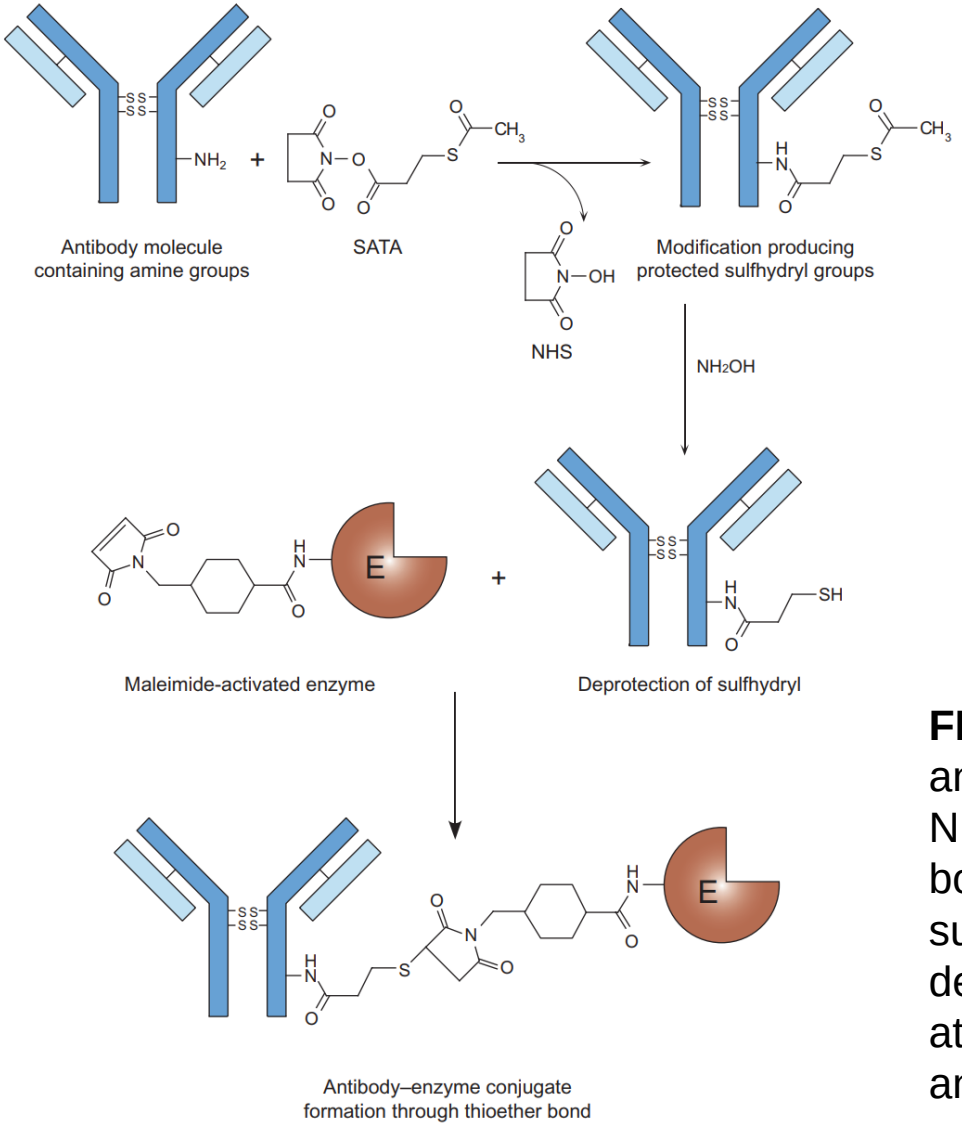
Antibody-Enzyme Conjugates

- HRP is incorporated in about 80% of all antibody conjugates, most of them utilized in diagnostic assay systems.
- AP is the second-most popular choice for antibody–enzyme conjugation, being used in almost 20% of all commercial enzyme-linked assays
- HRP is a glycoprotein and easily can be periodate oxidized for coupling via reductive amination to the amino groups on immunoglobulins

Antibody-Enzyme Conjugates



Antibody-Enzyme Conjugates



- The protected sulfhydryl group of SATA-modified proteins is stable to long-term storage without degradation

FIGURE 4.9 Available amine groups on an antibody molecule may be modified with the NHS ester end of SATA to produce amide bond derivatives containing terminal protected sulfhydryls. The acetylated thiols may be deprotected by treatment with hydroxylamine at alkaline pH. Reaction of the thiolated antibody with a maleimide-activated enzyme results in thioether crosslinks.

Antibody-Enzyme Conjugates

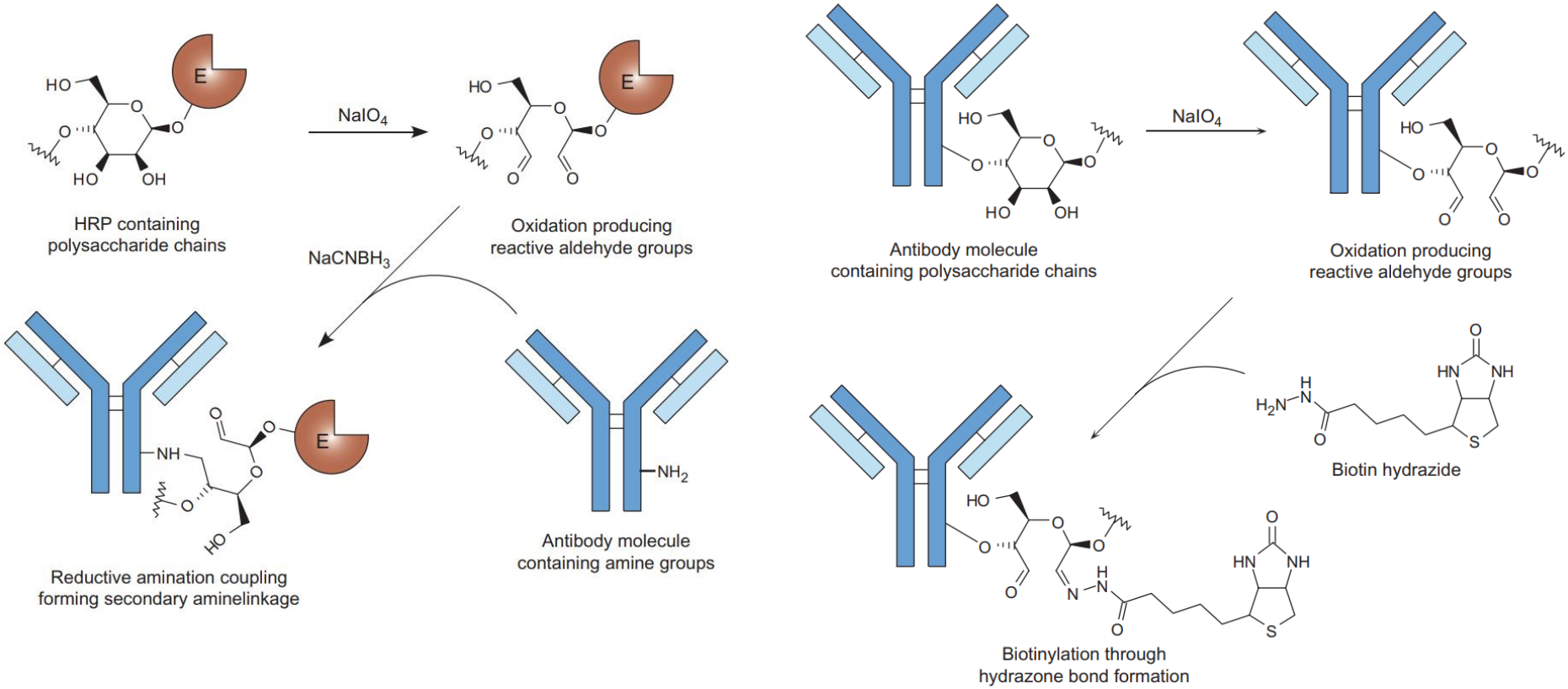
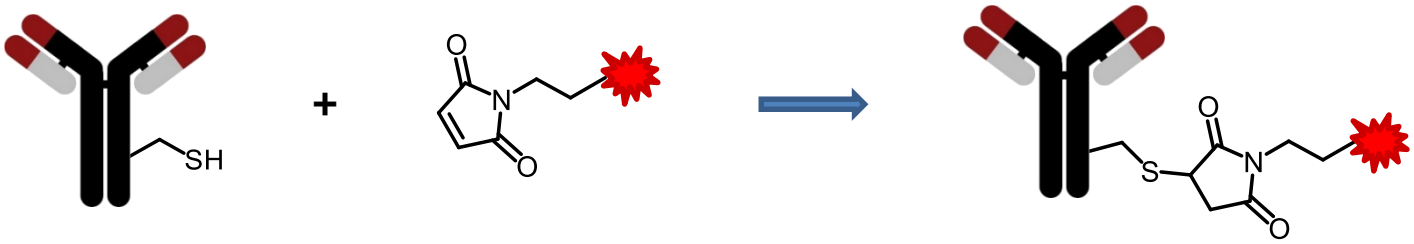
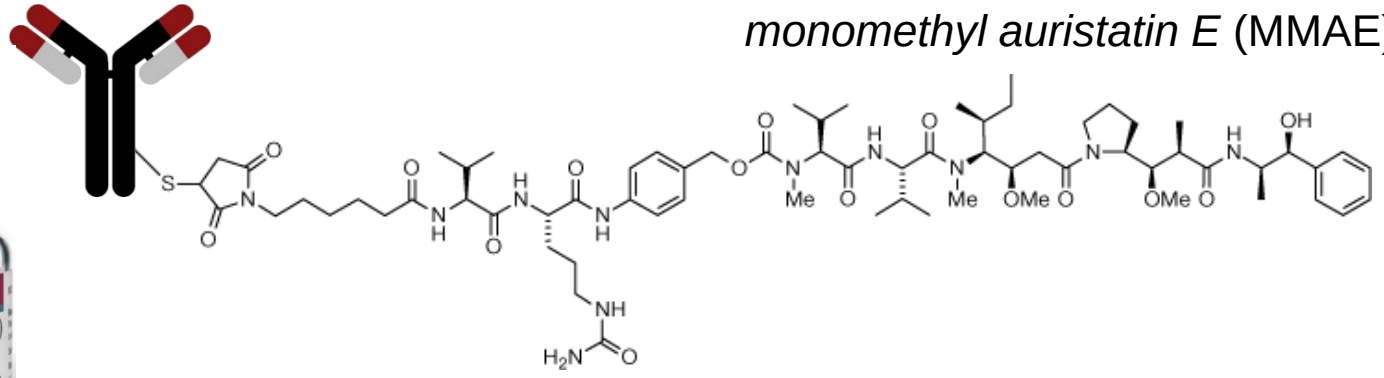


FIGURE 4.10 Polysaccharide groups on enzyme or antibody molecules may be oxidized with periodate to create aldehydes. Modification reactions can be carried out by using a reductive amination reaction or a imine formation reaction using a hydrazine or an alkoxyamine.

Antibody-Drug Conjugates

- *Brentuximab vedotin*, August 19, 2011, for relapsed Hodgkin lymphoma (HL) and relapsed systemic anaplastic large cell lymphoma (sALCL)



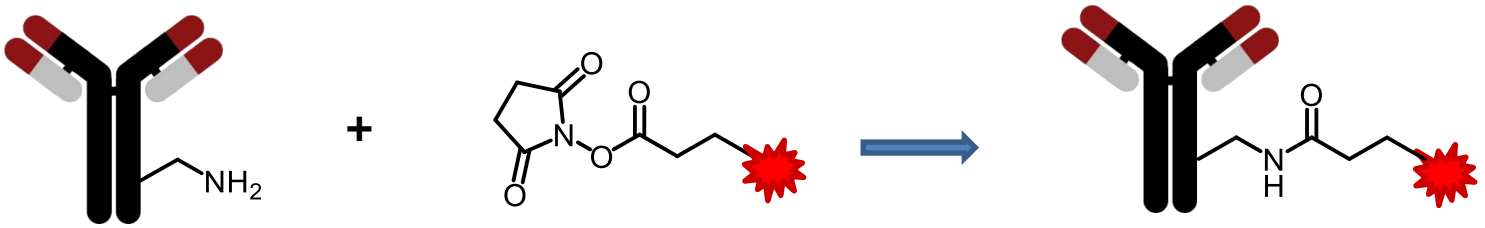


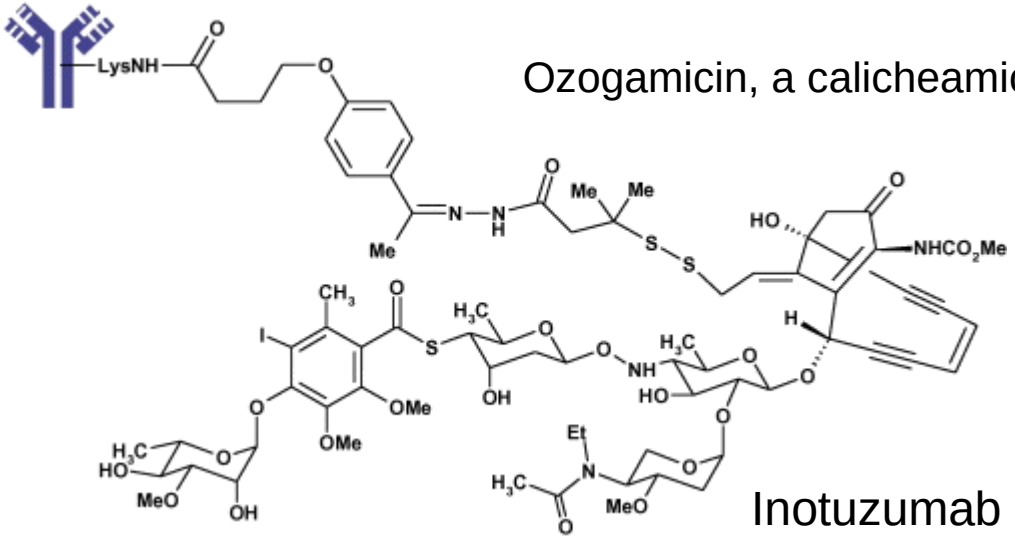
monomethyl auristatin E (MMAE)

Brentuximab vedotin

Antibody-Drug Conjugates

- *Inotuzumab ozogamicin,* June 30, 2017, for the treatment of adults with relapsed or refractory CD22-positive B-cell precursor acute lymphoblastic leukemia (ALL) under the trade name Besponsa®(Pfizer/Wyeth)



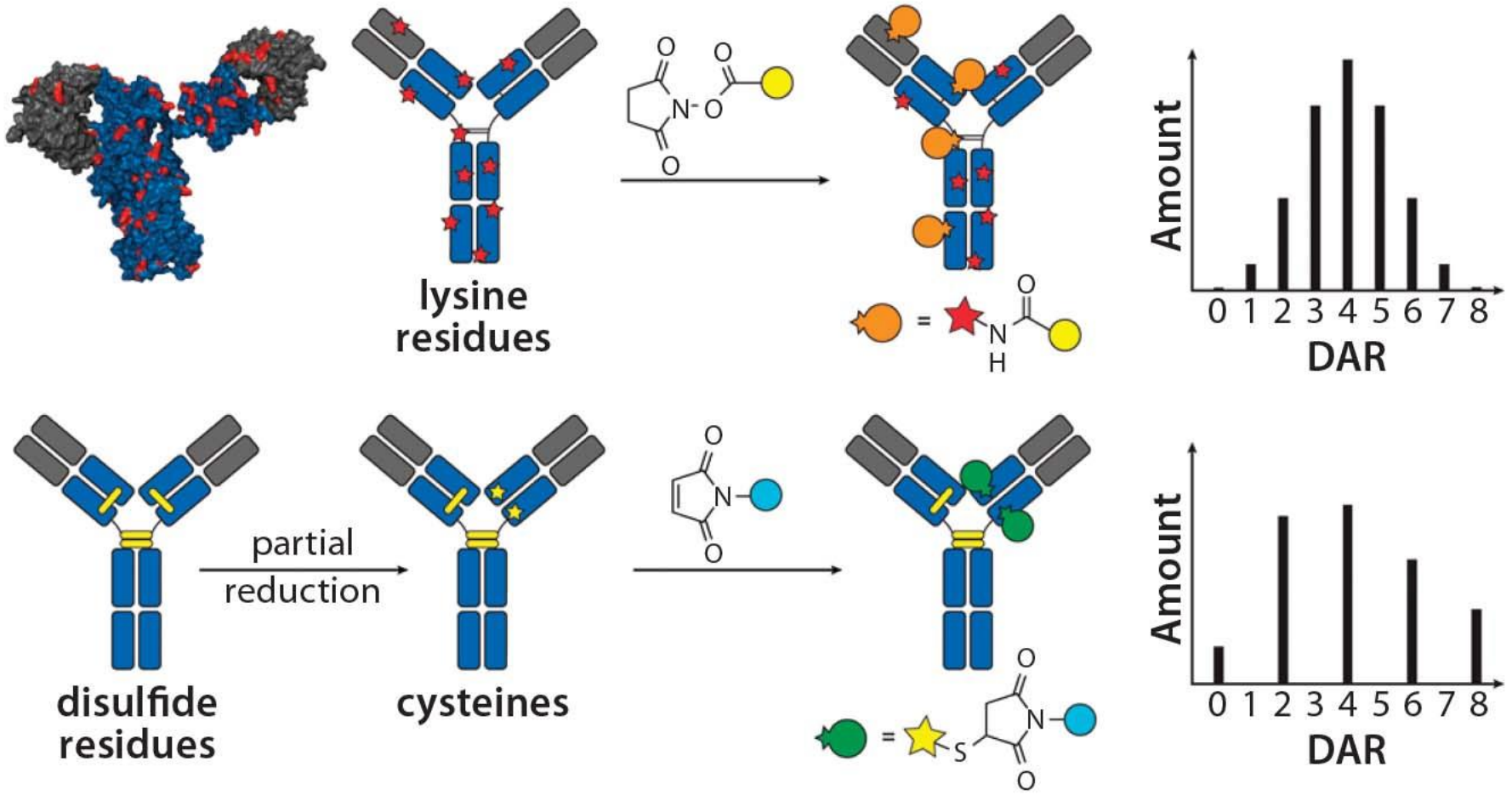


Ozogamicin, a calicheamicin derivative

Inotuzumab ozogamicin

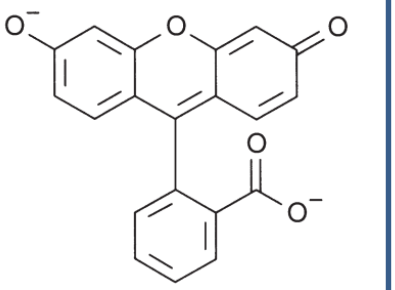
DAR 2~4

Drug-Antibody Ratios



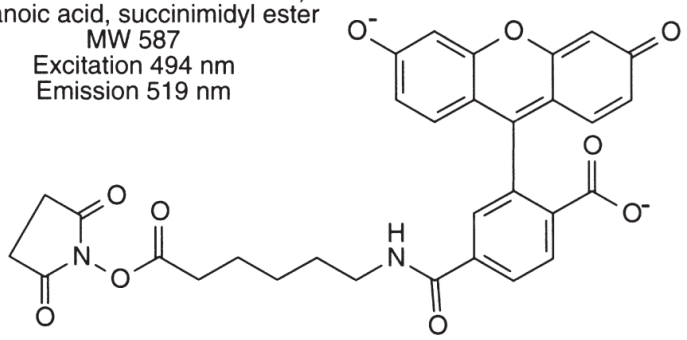
BioProcess International, **2014**.

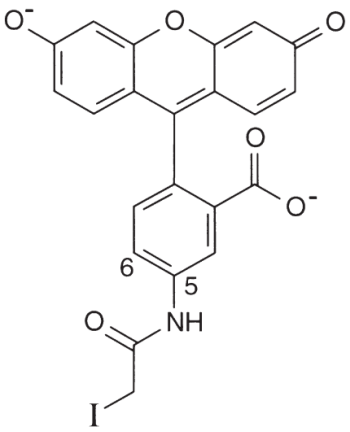
Fluoresceins and Rhodamines



Fluorescein

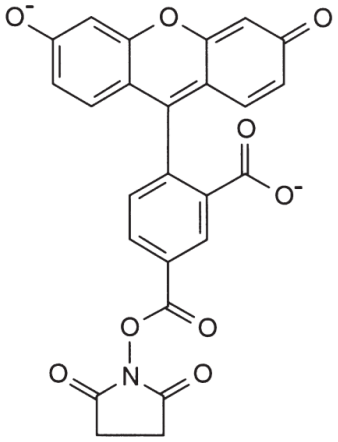
NHS-LC-Fluorescein;
6-(Fluorescein-5-carboxamido)
hexanoic acid, succinimidyl ester
MW 587
Excitation 494 nm
Emission 519 nm

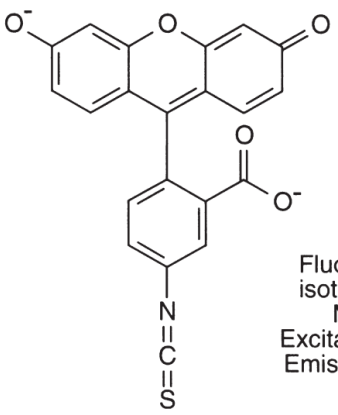




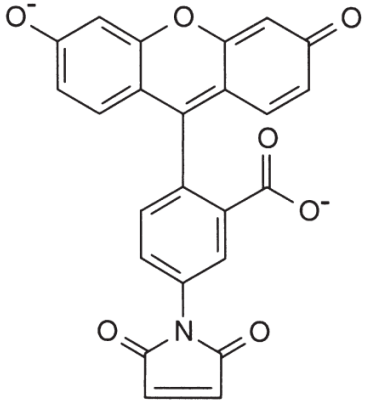
5- (and 6)-Iodoacetamido-Fluorescein
MW 515.26
Excitation: 490–495 nm
Emission: 515–520 nm
ε at 492 nm = 80,000–85,000 $M^{-1}cm^{-1}$

NHS-Fluorescein;
5-Carboxyfluorescein,
succinimidyl ester
MW 457
Excitation 491 nm
Emission 518 nm



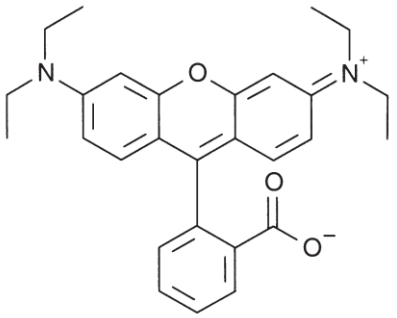


FITC;
Fluorescein-5-
isothiocyanate
MW 389
Excitation 494 nm
Emission 520 nm

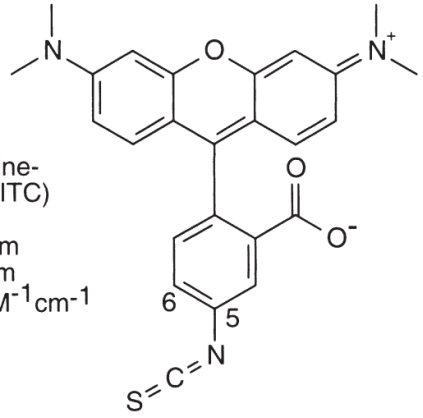


Fluorescein-5-maleimide
MW 427
Excitation: 490–495 nm
Emission: 515–520 nm
ε at 490 nm = 83,000 $M^{-1}cm^{-1}$

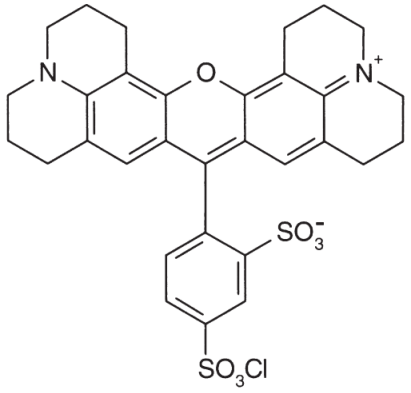
Fluoresceins and Rhodamines



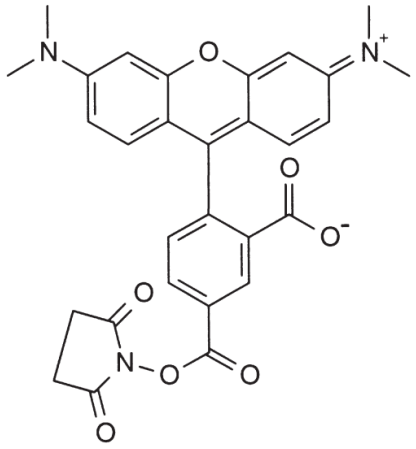
Rhodamine B



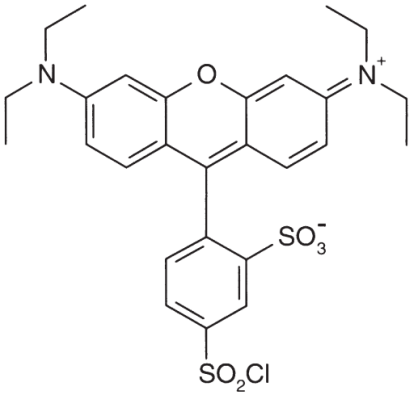
Tetramethylrhodamine-
5-isothiocyanate (TRITC)
MW 444
Excitation = 544 nm
Emission = 570 nm
ε at 544 nm = 100,000 $M^{-1}cm^{-1}$



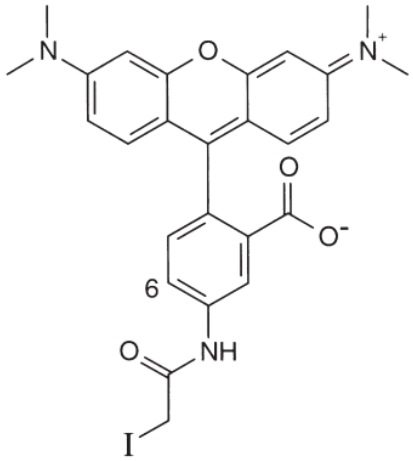
Texas Red Sulfonyl Chloride
MW 577
Excitation = 556 nm
Emission = 576 nm
ε at 556 nm = 93,000 $M^{-1}cm^{-1}$



NHS–Rhodamine;
5-Carboxytetramethyl-
rhodamine, succinimidyl ester
MW 528
Excitation = 546 nm
Emission = 579 nm
ε at 546 nm = 100,000 $M^{-1}cm^{-1}$

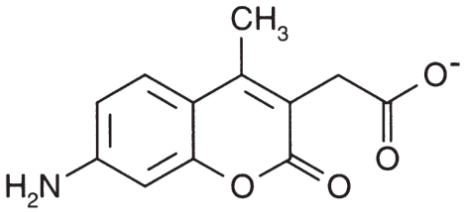


Lissamine Rhodamine B
Sulfonyl Chloride
MW 577
Excitation = 556 nm
Emission = 576 nm
ε at 556 nm = 100,000 $M^{-1}cm^{-1}$

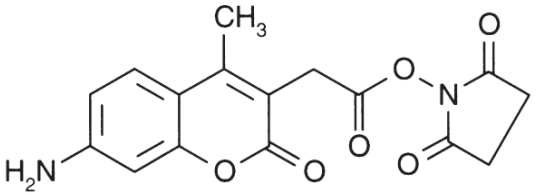


Tetramethylrhodamine-
5-iodoacetamide
MW 569
Excitation = 540 nm
Emission = 567 nm
ε at 540 nm = 76,000 $M^{-1}cm^{-1}$

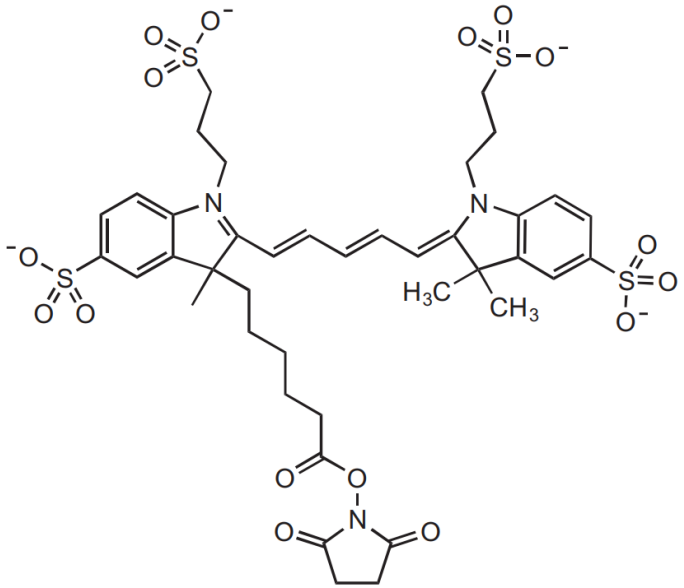
Other fluorophore derivatives



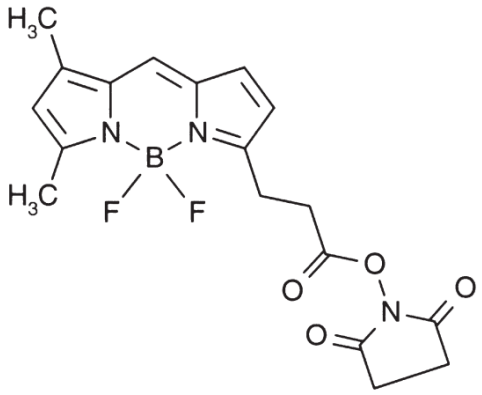
AMCA
7-Amino-4-methyl-
coumarin-3-acetic acid
MW 233
Excitation = 345–350 nm
Emission = 440–460 nm



AMCA–NHS



Tetrasulfonyl-Cy5-NHS ester



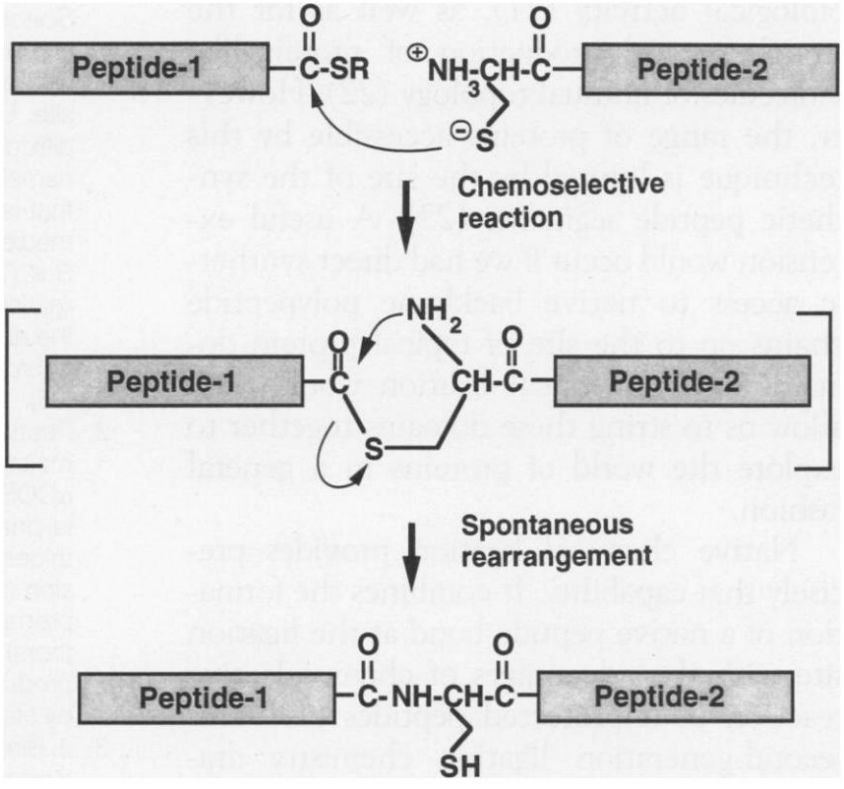
BODIPY FL C_3-SE
4,4-Difluoro-5,7-dimethyl-4-bora-3a,4a-
diaza-*s*-indacene-3-propionic acid,
succinimidyl ester
MW 389
Excitation = 502 nm
Emission = 510 nm
ε at 502 nm = 77,000 $M^{-1}cm^{-1}$

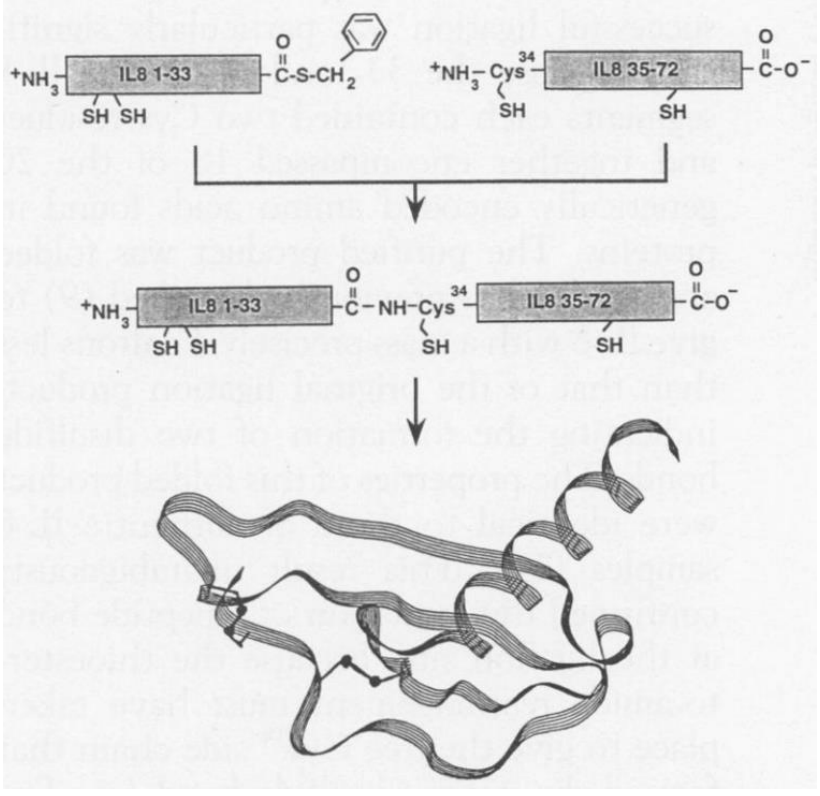
◆ Native Chemical Ligation

Synthesis of Proteins by Native Chemical Ligation

Philip E. Dawson, Tom W. Muir, Ian Clark-Lewis, Stephen B. H. Kent*

SCIENCE • VOL. 266 • 4 NOVEMBER 1994





Native Chemical Ligation

- Two peptides are synthesized by SPPS
- C-terminal thioester and N-terminal Cys
- Possible at physiological pH and in unprotected condition
- Slow reaction rate; a catalyst required
  4-carboxymethyl thiophenol
- Not applicable to a large protein

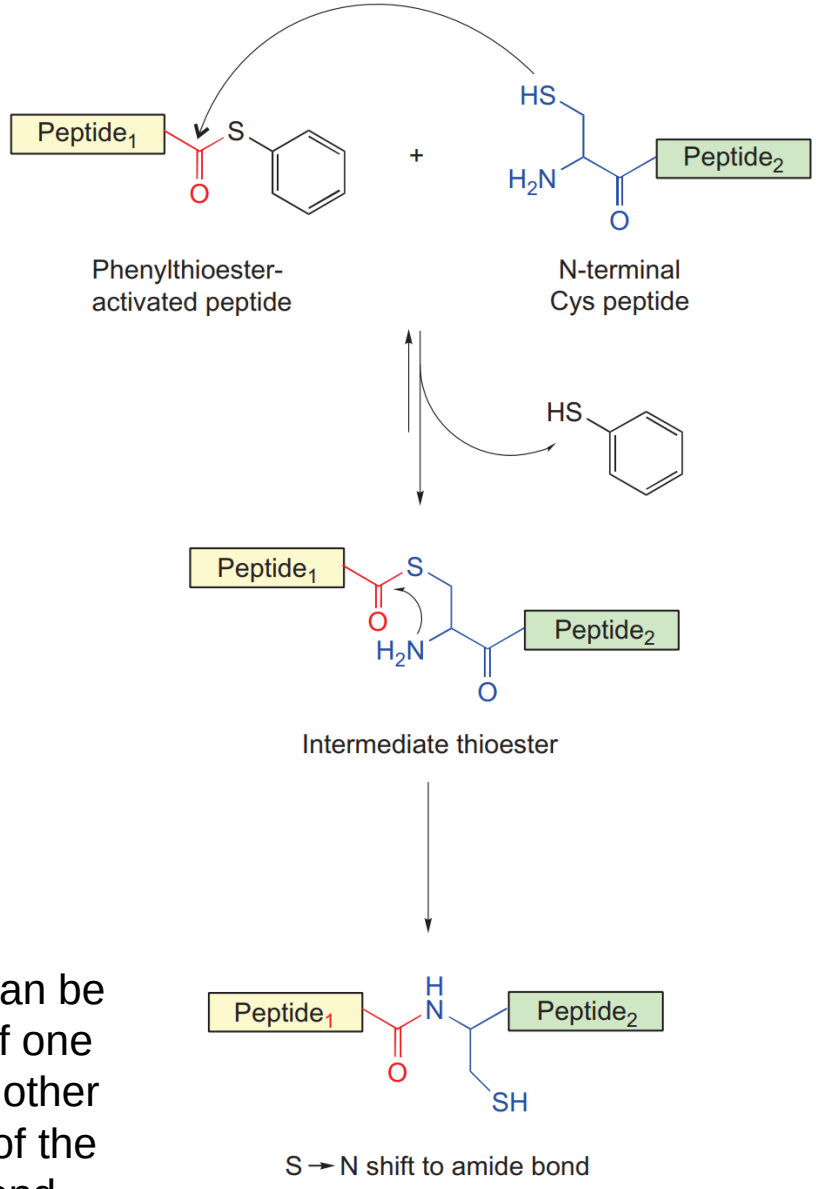


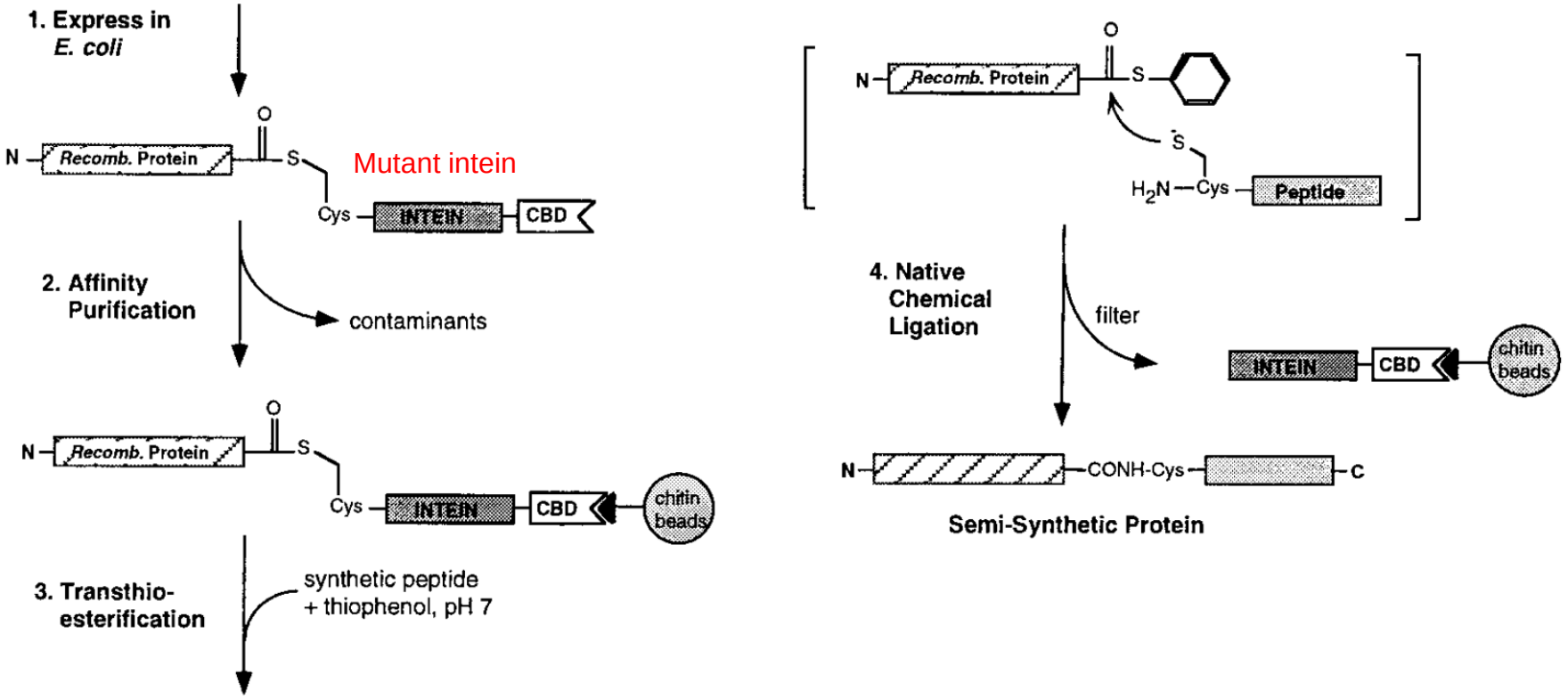
FIGURE 4.11 The native chemical ligation reaction can be used to form larger peptides from smaller peptides, if one contains a cysteine residue at its N-terminal and the other one contains a thioester on its C-terminal. Reaction of the peptide derivatives gives a native peptide (amide) bond.

Expressed Protein Ligation

Expressed protein ligation: A general method for protein engineering

TOM W. MUIR*†‡, DOLAN SONDHI†§, AND PHILIP A. COLE†‡§

Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 6705–6710, June 1998



Inteins

- An intein is a segment of a protein that is able to excise itself and join the remaining portions (the exteins) with a peptide bond during protein splicing
- The first intein was discovered in 1988 in yeast
- 113 known inteins are present in eukaryotes with minimum length of 138 amino acids and maximum length of 844 amino acids

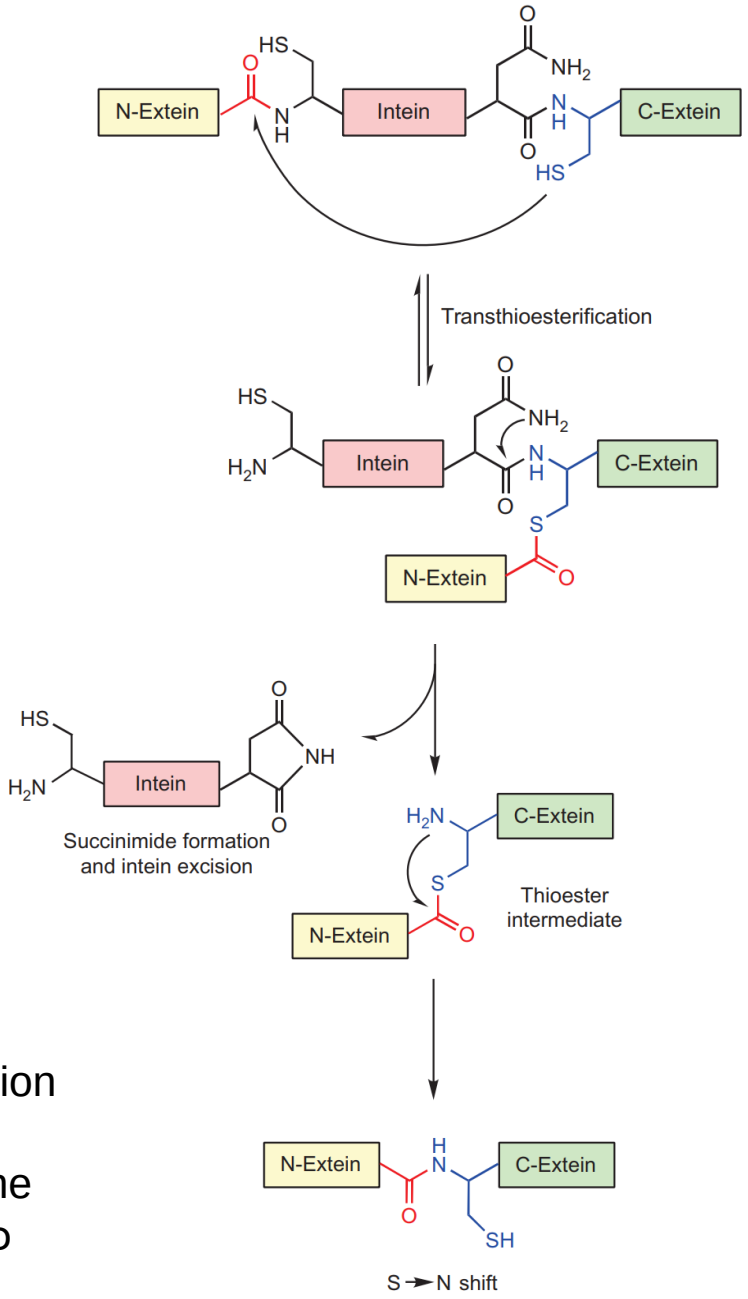
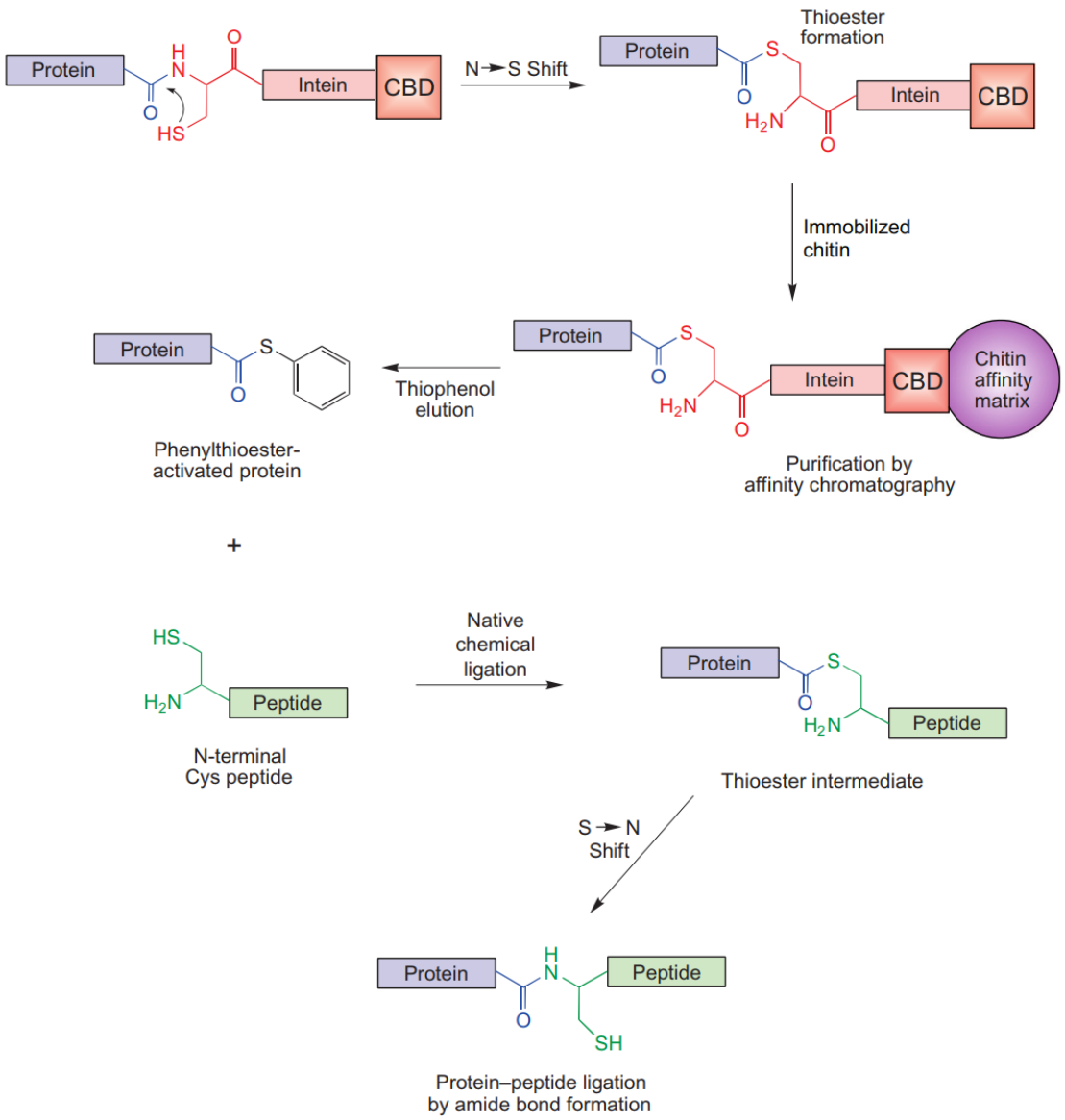


FIGURE 4.12 The native process leading to intein excision and ligation of extein fragments involves a sequence of reactions involving transthioesterification, cleavage of the intein fragment, and an S→N shift, which ligates the two extein peptides together via an amide bond.

Expressed Protein Ligation



- The expressed protein ligation process involves a fusion protein containing an intein tag plus a chitin binding domain
- The fusion protein is captured on an immobilized chitin resin and after removal of contaminating proteins is eluted using thiophenol
- This releases a phenylthioester-activated protein that can be used in the native chemical ligation reaction with another peptide containing an N-terminal cysteine residue
- Conjugation results in a native amide (peptide) bond formed between them.

Evaluation of chemical protein modification

- Efficiency
  - Quantitative and reproducible labeling efficiency
- Specificity
  - Labeling reactions need to occur at a desired location
  - Limitations in selecting a labeling location?
- General applicability
  - Generally applicable to various applications
  - Limitations in selecting a target protein?