이학박사 학위논문

New Bioorganic Strategies for the Generation of Functional Molecular Structures on Specific Residues in Peptides and Proteins

펩타이드 및 단백질 내 선택적 기능성 분자 구조 형성을 위한 신규 생체유기화학 전략

2020 년 2 월

서울대학교 대학원

화학부 생화학 전공

정 동 욱

Abstract

New Bioorganic Strategies for the Generation of Functional Molecular Structures on Specific Residues in Peptides and Proteins

Jung, Dongwook
Department of Chemistry, Biochemistry
The Graduate School
Seoul National University

Labeling biomolecules with functional chemical molecules plays a major role in understanding complex life phenomena and has led to the development of bioimaging and biotherapeutics. In particular, fluorescent substances among the functional molecules were tagged on peptides and proteins to allow for the observation of intracellular biomolecules behavior in the field of bio-imaging and for identifying the location and course of disease in real time by non-surgical methods. In addition, functional molecules having a therapeutic effect such as anticancer drugs can be effectively conjugated to antibodies, and polymers such as polyethylene glycol are bound to proteins to stabilize structures and improve biological activity.

The method widely used in the study of conjugating such functional molecules

to biomolecules is to stably bind the two substances through selective bioconjugation with natural amino acids present in the biomaterial and to introduce non-natural amino acids into specific sites through genetic code to give bioorthogonal functionalization. A widely used method for bioconjugation to natural amino acids is to chemically target amino acid residues to form bonds. In particular, the amine group of the lysine is often used through a nucleophilic reaction. In order to selectively bind to the amine of the lysine, various chemical functional groups should be introduced into the substance to be conjugated, and therefore, the amine is treated with *N*-hydroxysuccinimide, isothiocyanate, isocyanate and imidoester. In addition, many methods are used to ligate through the reaction of cysteine thiol and maleimide. Nevertheless, since most lysine is present on the surface of most proteins, various products with different activities are produced after the reaction, and since cysteine is easily oxidized and most are present in proteins in the form of disulfide bonds, the reduction process is necessary before the reaction is performed. Therefore, it is meaningful to develop a method for selectively conjugating biomolecules with molecules having various functionalities.

This doctoral dissertation involves 1) the formation of citrate-based fluorophore at the *N*-terminus of peptide and protein into *de novo*, and 2) the selective bioconjugation of tyrosine through sulfate-based reactions.

The method of forming the 5-oxo-2,3-dihydro-5H-[1,3]thiazolo[3,2-*a*] pyridine-3,7-dicarboxylic acid fluorophore is a dehydration and condensation reaction of citrate and cysteine. In this reaction, citrate and cysteine become amide bonds that require large amounts of energy. For this reason, it was confirmed that the

fluorophore was formed at the microwave method and at room temperature under peptide synthesis conditions by adding a coupling reagent that promotes amide bond formation. As a result, it was confirmed that a *de novo* fluorescent substance was formed under mild conditions using a citrate and a coupling reagent to introduce a cysteine-introduced peptide into the *N*-terminus. Furthermore, under a reaction condition through the use of a coupling reagent and citrate with *N*-terminal amino acids of proteins and peptides present in fixed cells and tissues, a fluorescent emission was confirmed by a confocal laser scanning microscope.

Sulfate click reactions proceed through the exchange of hexavalent sulfur and fluorine. Conventional sulfate click reactions are known as reactions between aromatic fluorosulfates and aryl silyl ethers, but no sulfate click reactions have been performed by directly activating an aromatic hydroxyl anion without a silyl group using a base. Therefore, in this study, we developed a sulfate click reaction condition so that only the tyrosine selectively reacts among various amino acids with a fluorosulfate moiety. Under these reaction conditions, only the tyrosine present in the TAT 47-57 cell penetrating peptide was selectively conjugated. Fluorescent peptides can be treated in cells and fluorescence can be confirmed in the cells using a confocal laser scanning microscope. Furthermore, the polyethylene glycol polymer was conjugated on the 49th tyrosine exposed on the surface of the erythropoietin known as an anemia drug by a sulfate click reaction. After that, the proportion of the volume of erythrocytes in the blood increases when the conjugated protein is intravenously administered *in vivo*.

In this study, the formation of selectively functional molecular structures in

peptides and proteins with new bioorganic strategies is expected to be widely used in the field of chemical biology as well as contributing to the fields of bioimaging and biotherapeutics.

Keywords: Bioconjugation, Fluorophore, *N*-terminal, *D*e *n*ovo, Coupling reagent, Sulfate click reaction, Tyrosine

Student Number: 2013-30907

Contents

Abstract ..... **i**
Contents ..... **v**
List of Figures ..... **viii**
List of Tables ..... **xv**

PART I. *De Novo* Formation of Citrate-based Fluorophores on *N*-Termini of Peptides and Proteins in Cells and Tissues
1. Abstract **1**
2. Introduction **2**
3. Experimental Section **4**
3.1 Materials 4
3.2. Instruments 5
3.3. Synthesis of 5-oxo-2,3-dihydro-5*H*-[1,3]thiazolo[3,2-*a*]pyridine-3,7-dicarboxylic acid (TPA) 5
3.4. Synthesis of Fmoc-Leu-Bu (2) 6
3.5. Synthesis of Leu-Bu (3) 6
3.6. Synthesis of Fmoc-Cys(Trt)-Leu-Bu (4) 7
3.7. Synthesis of Cys(Trt)-Leu-Bu (5) 7
3.8. Synthesis of Cys-Leu-Bu (6) 7
3.9. Synthesis of TPA-Leu-Bu (7) 8
3.10. Synthesis of TAT and LK peptides with an *N*-terminal cysteine on a solid phase peptide resin 8
3.11. Deprotection of S-(*tert*-butylthio) in TAT and LK peptides on the solid phase peptide resin 9
3.12. Formation of TPA on the peptides conjugated to the solid phase peptide resin 10
3.13. Cleavage of the TPA-peptides from the solid phase peptide resin 10
3.14. Cell culture 10
3.15. Cell fixation 11
3.16. Staining of fixed cells by the DNFC method 11

3.17. Labeling of intracellular organelles 12
3.18. CLSM observation of the DNFC-stained cells 12
3.19. Human tissue sample acquisition 12
3.20. Production of formalin fixed paraffin-embedded tissues and microsection 13
3.21. H&E- and MT-staining of human tissues for observation with optical microscopy 13
3.22. Staining of human tissues by the DNFC method 14
3.23. CLSM observation of the DNFC-stained tissues 14
4. Results and Discussion 15
4.1. Facile synthesis of Cit-based fluorophores at mild conditions 15
4.2. Fluorophore formation on *N*-termini of peptides 16
4.3. Fluorophore generation in cultured cell*s* 18
4.4. DNFC staining of human tissues 21
5. Conclusions 24
6. References 25

PART II. Chemoselective Tyrosine Bioconjugation Through Sulfate Click Reaction
1. Abstract 63
2. Introduction 64
3. Experimental Section 67
3.1 Materials 67
3.2. Instruments 67
3.3. Synthesis of phenyl *p*-tolyl sulfate (1) 68
3.4. Synthesis of *N*-Boc tyramine (2) 69
3.5. Synthesis of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (3) 69
3.6. Synthesis of tetraethylrhodamine succinimidyl ester (4) 69
3.7. Synthesis of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (5) 70
3.8. Synthesis of 4-((Fuorosulfonyl)oxy)benzoic acid (6) 70
3.9. Synthesis of 2,5-Dioxopyrrolidin-1-yl 4-((fluorosulfonyl)oxy) benzoate

(7) 71
3.10. Synthesis of methoxy poly ethylene glycol tosylate (8) 71
3.11. Synthesis of methoxy poly ethylene glycol azide (9) 72
3.12. Synthesis of methoxy poly ethylene glycol amine (10) 72
3.13. Synthesis of methoxy poly ethylene glycol amine (11) 72
3.14. Synthesis of (2S,2'S)-3,3'-((Sulfonylbis(oxy))bis(4,1-phenylene))bis(2-aminopropanoic acid) (12) 73
3.15. Synthesis of TAT 47-57 peptide 73
3.16. Synthesis of Rho-TAT peptide 74
3.17. Synthesis of PEG-rhEPO 75
3.18. Tryptic digestion of proteins 75
3.19. Confocal laser scanning microscopy (CLSM) observation of Rho-TAT peptide 75
3.20. *In vitro* cytotoxicity of diaryl sulfate compound 76
3.21. *In vivo* activity of rhEPO and PEG-rhEPO 77
4. Results and Discussion 78
4.1. Comparison of SuFEx reactivity of various model nucleophiles representing amino acids 78
4.2. The SuFEx application in TAT 47-57 peptide labeling by fluorophore 79
4.3. The SuFEx application in PEGylation of rhEPO 80
5. Conclusions 82
6. References 83

List of Publications 114
Abstract in Korean (국문 초록) 115

List of Figures

PART I. *De Novo* Formation of Citrate-based Fluorophores on *N*-Termini of Peptides and Proteins in Cells and Tissues

Figure 1. A scheme for formation of pyridone-based TPA and OPA structure by the reaction between Cit and Cys or Ser on proteins in cells and tissues.....28
Figure 2. A proposed mechanism of TPA formation without coupling reagents. (1) amide formation, (2) imide formation, (3) dehydration, (4) intramolecular condensation.....29
Figure 3. A proposed mechanism of TPA formation with PyBOP. (1): amide formation, (2): imide formation, (3): dehydration, (4): intramolecular condensation 29
Figure 4. Emission spectra (λ_{ex} = 365 nm) of mixtures containing the same amount of Cit and Cys after various treatments for 1 h: heating at 180°C (control); incubation with PyBOP at 25°C; incubation with PyBOP in presence of microwave irradiation. Each data point represents average ± S.D. (n = 4) 30
Figure 5. ^{1}H NMR spectrum (DMSO-d_6, 500 MHz) of TPA prepared by the PyBOP-assisted synthesis.....31
Figure 6. Absorption and emission spectra of TPA prepared by the PyBOP-assisted synthesis in deionized water at 25°C.....32
Figure 7. Emission spectra of mixtures in DMF containing the same amount of citric acid and L-cysteine which were treated with various reaction conditions.....33
Figure 8. (a) Emission spectra (λ_{ex} = 380 nm) of a mixture of citrate and L-serine in DMF for 1 h at 180°C (control) or incubation with PyBOP and DIPEA in presence of microwave irradiation for 1 h at 75°C. (b) Absorption and emission spectra (λ_{ex} = 380 nm) of a mixtures containing the same amount of citrate and L-serine in DMF after treatment in the same condition of (a) 34
Figure 9. Synthesis of TPA-Leu-Bu (**7**); model peptide in liquid phase 35
Figure 10. ^{1}H NMR spectrum (D_2O, 500 MHz) of Cys-Leu-Bu with a TPA structure (compound **7**) 36
Figure 11. Absorption and emission spectra of compound **7** in deionized water. Inset: Fluorescence of compound **7** illuminated by UV light 37

Figure 12. Synthetic scheme of TPA-peptides on SPPS resin. (1): deprotection of S-(*tert*-butylthio) group, (2): formation of TPA on the peptide, (3): cleavage of the TPA-peptide from the resin.....38
Figure 13. MALDI-TOF spectra of Cys-TAT, Cys-LK, TPA-TAT and TPA-LK peptides 39
Figure 14. (a) Absorption and emission spectra of TPA-TAT and TPA-LK peptides in deionized water. (b) PL intensity/Absorbance ratios of TPA, compound **7** and TPA-LK peptide in deionized water 40
Figure 15. CLSM images of *HeLa* cells treated with TPA-TAT and TPA-LK for 12 h.....41
Figure 16. Localization of TPA-labeled peptides in *HeLa* cells.....42
Figure 17. Simplified scheme of DNFC staining on cells.....43
Figure 18. Time-dependent development of fluorescence in *HeLa* cells stained by the DNFC method 44
Figure 19. (a) Emission spectra (λ_{ex}= 405 nm) of mixtures containing the same amount of citric acid and L-cysteine after incubation with PyBOP for 30 min at 25°C in various reaction solvents (black: DMF, red: DMSO, blue: DMSO/H_2O). (b) Comparison of fluorescence emission by excitation at the maximum absorption wavelength (maximum λ_{ex} of Cys/Cit= 370 nm and Ser/Cit= 400 nm) from a Cys/Cit mixture and a Ser/Cit mixture after incubation with PyBOP for 30 min at 25°C in DMSO/H_2O. (c) Comparison of fluorescence emission by excitation at 405 nm from a Cys/Cit mixture and a Ser/Cit mixture after incubation with PyBOP for 30 min at 25°C in DMSO/H_2O.....45
Figure 20. (a) CLSM images of *HeLa* cell samples stained by the DNFC method after 30 min in various solvents (solvents: DMF, DMSO and DMSO/H_2O). (b) CLSM images of the DNFC-stained *HeLa* cells at the focal plane near the plasma membrane (upper panel) and the magnified CLSM images of the white squares (lower panel). (c) CLSM images of the DNFC-stained *HeLa* cells using different laser sources. 46
Figure 21. Magnified CLSM images of *HeLa* cells stained by the DNFC method (N: nucleus, C: cytoplasm) 47
Figure 22. CLSM images of *HeLa* cells stained by the DNFC method and by each organelle-staining reagent. ER, mitochondria (Mito), and nuclear pore complex (NPC) were stained with concanavalin A (ConA), anti-TOMM20 antibody and anti-

NPC antibody, which are labeled with Alexa Fluor 594 (ConA and anti-NPC) or Alexa Fluor 647 (anti-TOMM20) 48
Figure 23. Colocalization ratio of blue/green fluorescence with red fluorescence. The colocalization degree was calculated from over 500 points in a CLSM image by the Imaris microscopy analysis software.....49
Figure 24. CLSM images of *HEK293T*, *THP-1M* and *NIH3T3* cells stained by the DNFC method. Without indication, the DNFC staining was performed for 30 min in DMSO/H_2O (70/30) 50
Figure 25. (a) CLSM images of a human skin/connective tissue and basal cell carcinoma (BCC) without the DNFC staining. The weak autofluorescence was observed in the image. (b) CLSM images of the DNFC-stained human skin tissue/connective tissue and BCC 51
Figure 26. Line profiles indicating the distribution of the blue (blue line) and green (green line) fluorescence in the fat tissue (a) without DNFC staining and (b) with DNFC staining through the white arrow line. (c) Comparison of the signal to noise (*S/N*) ratios of the fat tissues without DNFC staining (control; autofluorescence) and with DNFC staining 52
Figure 27. (a) CLSM images of a human skin/subcutaneous tissue sample strained with the DNFC method at a high magnification (H1: stratum spinosum compartment, arrows indicate polyhedral keratinocytes, H2: multiple layers of flake-like keratin in stratum corneum). (b) CLSM and optical microscopic (OM) images of skin/connective tissues stained by the DNFC and conventional histochemical methods (H&E and MT staining) at a low magnification 53
Figure 28. (a) CLSM and OM images of substructures of skin tissue samples, sebaceous glands (s. gland) and eccrine glands (e. gland), and basal cell carcinoma (BCC). (b) CLSM and OM images of substructures of subcutaneous tissues; arteriole and fat (arrows indicate histiocytes surrounding an adipocyte in fat necrosis)....54
Figure 29. ^{1}H NMR spectrum of TPA in DMSO-d_6 55
Figure 30. ^{1}H NMR spectrum of compound **2** in $CDCl_3$.....55
Figure 31. ^{1}H NMR spectrum of compound **3** in $CDCl_3$.....56
Figure 32. ^{1}H NMR spectrum of compound **4** in $CDCl_3$.....56
Figure 33. ^{1}H NMR spectrum of compound **5** in $CDCl_3$ 57
Figure 34. ^{1}H NMR spectrum of compound **6** in D_2O.....57
Figure 35. ^{1}H NMR spectrum of compound **7** in D_2O.....58

Figure 36. LC/MS spectra of TPA (a) and compound **7** (b) 59
Figure 37. MALDI-TOF spectra of Cys-TAT peptide (a), Cys-LK peptide (b), TPA-TAT peptide (c) and TPA-LK peptide (d).....60

PART II. Chemoselective Tyrosine Bioconjugation Through Sulfate Click Reaction

Figure 1. Artificial protein functionalization strategies 86
Figure 2. Comparison between SuFEx reaction developed by Sharpless group and stoichiometric DBU mediated desilylation reaction reported previously by us ... 86
Figure 3. A proposed mechanism for SuFEx reaction 87
Figure 4. Base screening in modified SuFEx reaction between *p*-cresol and phenyl fosylate 87
Figure 5. Reaction between phenyl fosylate and imidazole to yield a sulfamide product 88
Figure 6. ^{1}H NMR spectrum of phenyl *p*-tolyl sulfate (**1**) in $CDCl_3$ 89
Figure 7. ^{13}C NMR spectrum of phenyl *p*-tolyl sulfate (**1**) in $CDCl_3$ 89
Figure 8. ^{1}H NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$ 90
Figure 9. ^{13}C NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$ 90
Figure 10. ^{19}F NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$ 91
Figure 11. ^{1}H NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$ 92
Figure 12. ^{13}C NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$ 92
Figure 13. ^{19}F NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$ 93
Figure 14. ^{1}H NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$ 94
Figure 15. ^{13}C NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$ 94
Figure 16. ^{19}F NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$ 95

Figure 17. ^{1}H NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$ 96
Figure 18. ^{13}C NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$ 96
Figure 19. ^{19}F NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$ 97
Figure 20. ^{1}H NMR spectrum of methoxy poly ethylene glycol tosylate (**8**) in $CDCl_3$ 98
Figure 21. ^{1}H NMR spectrum of methoxy poly ethylene glycol amine (**10**) in $CDCl_3$ 98
Figure 22. ^{1}H NMR spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (**11**) in $CDCl_3$ 99
Figure 23. ^{19}F NMR spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (**11**) in $CDCl_3$ 99
Figure 24. ^{1}H NMR spectrum of (2S,2'S)-3,3'-((Sulfonylbis(oxy))bis(4,1-phenylene))bis(2-aminopropanoic acid) (**12**) in CD_3OD 100
Figure 25. GC spectrum of entry 1 in table 1 101
Figure 26. GC spectrum of entry 2 in table 1 101
Figure 27. GC spectrum of entry 3 in table 1 101
Figure 28. GC spectrum of entry 4 in table 1 102
Figure 29. GC spectrum of entry 5 in table 1 102
Figure 30. GC spectrum of entry 6 in table 1 102
Figure 31. GC spectrum of entry 7 in table 1 103
Figure 32. Functionalization of TAT 47–57 with a fluorescent small molecule 104
Figure 33. The HPLC chromatogram of TAT peptide (a) and Rho-TAT peptide (b). 104
Figure 34. MALDI-TOF MS spectra of TAT 47-57 and Rho-TAT 105
Figure 35. Large range MALDI-TOF MS spectrum of Rho-TAT peptide 105
Figure 36. MALDI-TOF MS spectra of reaction mixture of TAT 47-57 and Rho-Fs without TMG 106
Figure 37. MALDI-TOF/TOF spectra of TAT 47-57 (upper) and Rho-TAT 106
Figure 38. CLSM images of *HeLa* cells after 12 h-treatment of deionized water (a-d), and Rho-TAT peptide (e-h) 107
Figure 39. The schematic representation of the SuFEx reaction between rhEPO and

a PEGylating reagent.....108
Figure 40. MALDI-TOF MS spectra of rhEPO and PEG-rhEPO.....108
Figure 41. MALDI-TOF MS spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (**11**) 109
Figure 42. MALDI-TOF MS spectra of rhEPO (upper) and PEG-rhEPO after trypsin treatment.....109
Figure 43. The HCT profiles of control, rhEPO-treated, PEG-rhEPO-treated Balb/c mice.....110
Figure 44. Dose dependent cytotoxicity of **12** in *HeLa* cells for 24 h (a) and 48 h (b) 111

List of Tables

PART I. *De Novo* Formation of Citrate-based Fluorophores on *N*-Termini of Peptides and Proteins in Cells and Tissues

Table 1. Summary of DNFC-staining characteristics of cells in this research 61
Table 2. Summary of DNFC-staining characteristics of tissues in this research .62

PART II. Chemoselective Tyrosine Bioconjugation Through Sulfate Click Reaction

Table 1. Comparison of SuFEx reactivity in the reactions of various model nucleophiles representing amino acids with phenyl fosylate 112
Table 2. PEGylated VNFYAWKR *m/z* peaks in MALDI-TOF spectrum of PEG-rhEPO after trypsin treatment 113
Table 3. PEGylated VNFYAWK *m/z* peaks in MALDI-TOF spectrum of PEG-rhEPO after trypsin treatment 113

PART I. *De Novo* Formation of Citrate-based Fluorophores on *N*-Termini of Peptides and Proteins in Cells and Tissues

1. Abstract

We developed a new method for the *de novo* formation of thiazolopyridone- or oxazolopyridone-based fluorophores from non-fluorogenic citrate and cysteine or serine under mild reaction conditions. Use of an amide coupling reagent and microwave irradiation greatly facilitates the fluorophore formation on peptides and proteins with *N*-terminal cysteine or serine residues in biosamples. The fluorophore was introduced on the *N*-terminal cysteine peptides during solid phase peptide synthesis (SPPS) without disruption of the peptide function. Since *N*-terminal cysteine and serine in proteins can form thiazolopyridone- or oxazolopyridone-based fluorophores emitting blue and green fluorescence, respectively, in the condition of the *de novo* fluorophore formation based on citrate (DNFC), each organelle, cell and tissue exhibited a characteristic blue and green fluorescence distribution by the DNFC staining. We believe that the DNFC staining can provide a new potential protocol for future cell imaging, histology and diagnosis.

2. Introduction

Fluorescence-based imaging is a current paradigm method for visualising biological samples as it is highly sensitive, has a low detection limit and is non-invasive.[1, 2] Various fluorescent species have been developed for obtaining desired excitation/emission wavelengths, improved quantum yields and negligible photobleaching. Well-designed chemical fluorophores, such as organic dyes,[3, 4] inorganic particles[5, 6] and their hybrids[7-9] are commonly conjugated or complexed to biological molecules and generate corresponding fluorescence signals on specific targets in biosystems. Unlike most chemical fluorophores having fluorogenic structures that were formed prior to their introduction into biological molecules, some biological fluorophores are gradually formed via maturation from almost non-fluorogenic structures. Fluorophore formation in green fluorescent protein from threonine, tyrosine and glycine in the amino acid sequence is a representative example.[10, 11] This research on *de novo* fluorophore generation will improve understanding of the origin of biological fluorophores as well as facilitate the discovery of new ones, although few examples have been reported so far except for biological fluorophores.

In a previous study, Yang's group reported that fluorescence properties could be induced in polymers by heating with citrate (Cit) and different non-fluorogenic amino acids.[12] The strongest fluorescence occurred when L-cysteine (Cys) reacted with Cit. Kasprzyk's group subsequently confirmed that the fluorescence originated from the structure of 5-oxo-2,3-dihydro-5H-[1,3]thiazolo[3,2-*a*]pyridine-3,7-dicarboxylic acid (TPA).[13] It was proposed that the dihydrothiazolopyridone ring in TPA is formed via serial condensation/dehydration reactions between carboxylic acid and amine/thiol groups in Cit and Cys, respectively. As the condensation/dehydration reactions normally require extreme conditions without catalysts, the reported syntheses proceeded at high temperatures (140°C–180°C).[13-

15] Furthermore, the extreme reaction conditions frequently led to heterogeneous product mixtures of oligomers and polymers of TPA or even carbon dots with varying luminescence properties.[16, 17]

The *de novo* TPA fluorophore formation from natural non-fluorogenic tricarboxylic acid and amino acid may be extremely useful for introducing fluorescence into native peptides and proteins if the fluorophore can be generated under mild conditions (Figure 1). In this study, we intended to develop a method to facilitate formation of the Cit-based fluorophores at reaction conditions avoiding significant disruption of biomolecular structures. Then these conditions were applied to introduce the fluorophores into synthetic peptides and native proteins in cells. We carefully analysed the fluorescence distribution among subcellular structures to understand the fluorophore generation on cellular proteins. Finally, we examined the potential of the staining based on the de novo fluorophore formation based on citrate (DNFC) as a new candidate for the histological analysis by comparing the characteristic fluorescence generation in substructures in human tissues with conventional histochemical staining.

3. Experimental Section

3.1. Materials

N-α-Fmoc protected L-amino acids, Rink Amide MBHA resin (0.45 mmol/g loading), and benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) were purchased from BeadTech, Korea. Citric acid monohydrate was purchased from DAEJUNG, Korea. L-cysteine hydrochloride monohydrate (L-Cys·HCl·H_2O) was purchased from BIOSESANG INC., Korea. Glycine was purchased from Merck Millipore, USA. *N*,*N*-Diisopropylethylamine (DIPEA), triisopropylsilane (TIS), glutaraldehyde (GA), and 1,2-ethanedithiol (EDT) were purchased from Tokyo Chemical Industry (TCI), Japan. Dimethyl sulfoxide (DMSO), n-butylamine, *N*,*N*′-diisopropylcarbodiimide (DIC), L-serine, L-methionine, 1-hydroxybenzotriazole hydrate (HOBt), DL-dithiothreitol (DTT), trifluoroacetic acid (TFA), triethylsilane (TES), paraformaldehyde (PFA), gelatin, sodium borohydride and piperidine were purchased from Sigma-Aldrich, USA. *N*,*N*-Dimethylformamide (DMF), 1,2-dichloromethane (DCM), acetonitrile (ACN), ethyl acetate (EtOAc), *n*-hexane (*n*-Hex) and diethyl ether were purchased from Samchun Chemical, Korea. *HeLa* (human cervical cancer), *THP-1* (human monocyte cell) and *NIH3T3* (mouse fibroblast) cells were purchased from Korean Cell Line Bank. *HEK293T* (human embryonic kidney) cells were purchased from American Type Culture Collection (ATCC). SYTO 59 red fluorescent nucleic acid stain (5 mM solution in DMSO) and concanavalin A, Alexa Fluor 594 conjugate were purchased from Thermo Fisher Scientific, USA. Anti-TOMM20 antibody [EPR15581-54] - Mitochondrial Marker (Alexa Fluor® 647), Anti-NUP133 antibody [EPR10808(B)] and Donkey Anti-Rabbit IgG H&L (Alexa Fluor® 594) were purchased from abcam. Dulbecco's modified eagle's medium (DMEM), Dulbecco's phosphate buffered saline (DPBS) and fetal bovine serum (FBS) were purchased from WELGENE,

Korea. 0.05% Trypsin-EDTA (1×) was purchased from Gibco, USA.

3.2. Instruments

All ^{1}H NMR-spectra were analyzed with a Varian NMR (500 MHz) from USA and a Bruker Avance DPX-300 (300 MHz) from Germany. Microwave irradiation for synthesis and modification of peptides was provided by a CEM Discover microwave from USA. Photoluminescence excitation and emission were measured with a JASCO FP-8300 spectrofluorometer from Japan. High-performance liquid chromatography (HPLC) analyses and separation were carried out on a LC-20 series SHIMADZU system (Japan) equipped with a C18 Zorbax (5 μm, 9.4 × 250 mm) at 30^{o}C. 0.1% TFA in water and acetonitrile were used a as eluents at a flow rate of 4 mL/min. Quality of peptides was monitored with a MALDI-TOF mass analyzer (Bruker, Germany, DE/microflex LT). LC-ESI-MS measurements were performed on an Agilent 6100 series system equipped with an Agilent 6120 quadrupole MS. Confocal laser scanning microscopic (CLSM) images were acquired using a Zeiss DE/LSM 510 NLO (Carl Zeiss, Germany) with a 500× objective lens (C-Apochromat, Carl Zeiss).

3.3. Synthesis of 5-oxo-2,3-dihydro-5*H*-[1,3]thiazolo[3,2-*a*]pyridine-3,7-dicarboxylic acid (TPA)

Citric acid monohydrate (0.100 g, 0.476 mmol) was dissolved in DMF (0.5 mL) and DIPEA (81.0 μL, 0.476 mmol) was added to the solution. PyBOP (0.248 g, 0.476 mmol) in DMF (1 mL) was added to the citric acid solution. L-Cys · HCl · H_2O (0.085 g, 0.48 mmol) in DMF (0.5 mL) was added to the yellowish solution. Then, the solution was heated by microwave at 75^{o}C (25 W) or stirred at room temperature (25^{o}C). After the reaction, the product with a strong absorption band at 365 nm was

purified using HPLC. The fraction was collected and freeze-dried to obtain the pale yellowish solid. ^{1}H NMR (500 MHz, DMSO-d_6): δ 6.64-6.60 (s, 1H), 6.55-6.51 (s, 1H), 5.51-5.42 (d, 1H), 3.97-3.85 (m, 1H), 3.66-3.56 (m, 1H). LC/MS (*m/z*): $[M+H]^+$ calcd. for $C_9H_7NO_5S$; 241.0 found. 242.2 $[M+H]^+$.

3.4. Synthesis of Fmoc-Leu-Bu (2)

Fmoc-Leu-OH **(1)** (1.00 g, 2.83 mmol) and DIC (0.438 mL, 2.83 mmol) were dissolved in DCM (10 mL). Then, HOBt (0.382 g, 2.83 mmol) dissolved in DCM (10 mL) and n-butylamine (0.336 mL, 3.34 mmol) were added to the **(1)** solution. The solution was stirred at room temperature for 3 h. The reaction solution was concentrated via rotary evaporation. The mixture was purified by silica gel chromatography (*n*-Hex:EtOAc=1:1) to obtain a white solid (yield 87%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.78-7.28 (m, 8H), 4.43-4.41 (d, 2H), 4.23-4.19 (t, 1H), 4.12 (s, 1H), 3.25-3.23 (d, 2H), 1.61-1.25 (m, 6H), 0.93-0.88 (m, 9H).

3.5. Synthesis of Leu-Bu (3)

Compound **(2)** (0.100 g, 0.245 mmol) was dissolved in 10% piperidine in DCM (10 mL). The solution was stirred at room temperature for 1 h. The reaction solution was concentrated via rotary evaporation. The product was obtained as a white solid (yield 71%) by silica gel chromatography (DCM:MeOH=10:1). ^{1}H NMR (300 MHz, DMSO-d_6): δ 3.10-3.04 (m, 3H),1.74-1.61 (m, 1H), 1.42-1.14 (m, 6H), 0.89-0.82 (m, 9H).

3.6. Synthesis of Fmoc-Cys(Trt)-Leu-Bu (4)

Compound **(3)** (0.100 g, 0.537 mmol), HOBt (72.6 mg, 0.537 mmol), and Fmoc-Cys(Trt)-OH(315 mg, 0.537 mmol) were dissolved in DCM (2 mL). Then, DIC (83.2 μL, 0.537 mmol) was added to the **(3)** solution. The solution was stirred at room temperature for 3 h. The reaction solution was concentrated via rotary evaporation. The product was obtained as a white solid (yield 24%) by silica gel chromatography (*n*-Hex:EtOAc=3:1). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.18-7.19 (m, 23H), 4.89-4.87 (d, 1H), 4.41-4.29 (m, 3H), 4.20-4.15 (t, 3H), 3.13-3.11 (d, 2H), 2.73-2.59 (m, 2H), 1.78-1.75 (m, 1H), 1.58-1.20 (m, 6H), 0.89-0.82 (m, 9H).

3.7. Synthesis of Cys(Trt)-Leu-Bu (5)

Compound **(4)** (0.100 g, 0.133 mmol) was dissolved in 10% piperidine in DCM (1 mL). The solution was stirred at room temperature for 1 h. The reaction solution was concentrated via rotary evaporation. The product was obtained as a pale yellowish solid (yield 36%) by silica gel chromatography (*n*-Hex:EtOAc=5:1 and then, EtOAc only). ^{1}H NMR (300 MHz, $CDCl_3$): δ 7.45-7.19 (m, 15H), 4.28-4.21 (m, 1H), 3.26-3.08 (m, 2H), 3.03-2.99 (q, 1H), 2.69-2.50 (m, 2H), 1.69-1.32 (m, 7H), 0.91-0.85 (m, 9H).

3.8. Synthesis of Cys-Leu-Bu (6)

Compound **(5)** (0.100 g, 0.188 mmol) was dissolved in DCM (1 mL). TES (0.300 mL, 1.88 mmol) and TFA (3 mL) were added to the **(5)** solution. The solution was stirred at room temperature for 2 h. After the reaction, the solution was concentrated by a nitrogen gas blowing and rotary evaporation. Then, a white solid

was obtained by precipitaion in diethyl ether. The precipitate was collected by centrifugation at 4,000 rpm for 10 min and washed with diethyl ether (×4). The white solid product was dried by high vacuum (yield 31%). ^{1}H NMR (300 MHz, D_2O): δ 4.34-4.23 (m, 2H), 3.33-3.02 (m, 4H), 1.68-1.44 (m, 5H),1.36-1.24 (m, 2H), 0.95-0.86 (m, 9H).

3.9. Synthesis of TPA-Leu-Bu (7)

Citric acid monohydrate (5.0 mg, 0.024 mmol) was dissolved in DMSO (50 μL) and DIPEA (4.1 μL, 0.024 mmol) was added to the solution. PyBOP (12.4 mg, 0.0240 mmol) in DMSO (100 μL) was added to the citric acid solution. Compound **(6)** (6.9 mg, 0.024 mmol) was dissolved in DMSO (50 μL) and added to the citric acid solution. Then, the solution was heated by microwave at 75°C (25 W) for 30 min. After the reaction, the product with a strong absorption at 365 nm was purified using HPLC. The purified fraction was freeze-dried to obtain a pale yellowish solid (yield 13%). ^{1}H NMR (500 MHz, D_2O): δ 6.75 (s, 1H), 6.59 (s, 1H), 5.61-5.60(d, 1H), 4.33-4.32(m, 1H), 4.00-3.96(m, 1H), 3.57-3.53(m, 1H), 3.25-3.15(m, 2H), 1.68-1.29(m, 7H), 0.96-0.88(m, 9H). LC/MS (*m/z*): calcd. for $C_{19}H_{27}N_3O_5S$; 409.2 found. 410.2 $[M+H]^+$.

3.10. Synthesis of TAT and LK peptides with an *N*-terminal cysteine on a solid phase resin

TAT and LK peptides with an *N*-terminal cysteine (sequence: CYGRKKRRQRRR and CLKKLLKLLKKLLKLAG) was synthesized by Fmoc-based solid-phase peptide synthesis chemistry using a Rink Amide MBHA resin. The peptide was synthesized in a 90 μmol scale. The typical conditions for the

microwave-assisted peptide synthesis were used:[18, 19] Fmoc-deprotection with 20% piperidine at 75ºC (50 W) for 20 min and peptide coupling with PyBOP and DIPEA at 75ºC (25 W) for 20 min. First of all, Rink Amide resins (200 mg, 0.090 mmol, 0.45 mmol/g loading) were deprotected with 20% peperidine in DMF (2 mL) under N_2 bubbling, and then the first amino acid (350 mg, 0.540 mmol), PyBOP (281 mg, 0.540 mmol) and DIPEA (200 μL, 1.18 mmol) were added to the suspension. With microwave irradiation, the suspension was stirred under N_2 atmosphere for 20 min. Reaction byproducts were removed with DMF and DCM several times. After the coupling step, Fmoc-protected amino acid was deprotected with 20% piperidine in DMF and the reaction byproducts were removed with DMF and DCM several times. The coupling and deprotection steps were repeated with different Fmoc-protected amino acids sequentially until the deprotection of the last Fmoc-protected amino acid (Fmoc-Cys) was complete.

3.11. Deprotection of S-(*tert*-butylthio) in TAT and LK peptides on the solid phase peptide resin

For deprotection of S-(*tert*-butylthio) protecting group with an *N*-terminal cysteine amino acid in the prepared peptides on the resin, the swollen resin (100 mg, 0.045 mmol) was reacted with DTT (250 mg, 1.62 mmol) in DMF (2 mL) for 1 h at 75ºC (50 W) by using microwave heating. Reaction byproducts were removed with DMF and DCM several times. LRMS (MALDI) *m/z*: Anal. calcd. for $[M+H]^+$ $C_{67}H_{123}N_{33}O_{15}S$ (Cys-TAT): 1662.97; found: 1662.50. Anal. calcd. for $[M+H]^+$ and $[M+Na]^+$ $C_{92}H_{176}N_{24}O_{17}S$ (Cys-LK): 1923.35, 1945.33; found: 1923.27, 1945.17.

3.12. Formation of TPA on the peptides conjugated to the solid phase peptide resin

Citric acid monohydrate (189 mg, 0.900 mmol) was dissolved in DMF (1 mL). Then, DIPEA (154.4 μL, 0.90 mmol) and PyBOP (468 mg, 0.90 mmol) dissolved in DMF (1 mL) was added to the citric acid solution. The swollen resin (100 mg, 0.045 mmol) in DMF (2 mL) were reacted with the prepared solution and EDT (50 μL) for 1 h at 75°C (25 W) by using microwave heating.

3.13. Cleavage of the TPA-peptides from the solid phase peptide resin

Cleavage of the peptide from the resin was carried out by treatment with TFA/TIS/water (total volume 4.00 mL, v/v=95:2.5:2.5) for 4 h at room temperature. The resin was then separated by filtration and further washed with TFA. The filtrate was concentrated by a nitrogen gas blowing. The synthesized peptide was precipitated in a mixture of *n*-hexane and diethyl ether (v/v=50:50). The resulting suspension was centrifuged at 4,000 rpm for 15 min. After the supernatant was decanted, the pellet was dissolved in DMF and was purified with HPLC. The peptide was lyophilized affording a white powder. LRMS (MALDI) *m/z*: Anal. calcd. for $[M+H]^+$ $C_{73}H_{124}N_{33}O_{18}S$ (TPA-TAT): 1782.94; found: 1782.87 and $C_{98}H_{177}N_{24}O_{20}S$ (TPA-LK): 2043.30; found: 2044.87.

3.14. Cell culture

HeLa (human cervical cancer), *HEK293T* (human embryonic kidney) and *NIH3T3* (mouse fibroblast) cells were cultured in DMEM supplemented 10% FBS. *THP-1* (human monocyte cell) cells were cultured in RPMI 16409 supplemented 5%

FBS. The *THP-1* mococytes were differentiated into macrophages (*THP-1M*) by treatment of with phorbol 12-myristate 13-acetate (PMA). *THP-1* cells (50,000 cells/well) were seeded on gelatin coated coverslips and incubated with PMA (10 ng/mL) for 72 h. All the cells were maintained in a humidified atmosphere of 5% CO_2 at 37°C in a incubator.

3.15. Cell fixation

14-mm^2 glass coverslips were coated with 0.1% gelatin in ultrapure water. Coverslips were placed in a 6-well plate and seeded with 50,000 cells overnight. For ER and mitochondria staining, cells were washed and fixed with 3% PFA + 0.1% GA in PBS for 10 min, followed by reduction with 0.1% sodium borohydride for 7 min at RT and quenching in 100 mM glycine for 10 min at RT. Finally, cells were washed with PBS three times for 5 min each. For nuclear pore complex (NPC) staining, cells were rinsed with 2.4% PFA in PBS and extracted with 0.4% Triton X-100 in PBS for 3 min. And the cells were fixed with 2.4% PFA in PBS for 30 min, quenched for 5 min with 50 mM NH_4Cl solution, and washed with PBS three times for 5 min each.

3.16. Staining of fixed cells by the DNFC method

First, Citric acid monohydrate (0.010 g, 0.0476 mmol) was dissolved in DMSO/H_2O (v/v=70/30, 0.5 mL) and DIPEA (8.1 μL, 0.0476 mmol) was added to the solution. Then PyBOP (0.0248 g, 0.0476 mmol) was dissolved in DMSO/H_2O (v/v=70/30, 1.0 mL) and the solution was added to the citric acid solution. Fixed cells in glass chamber were stained with the prepared solution at room temperature. Generally, the staining was performed under the dark and shaking for 30 min. After

the staining, the cells were washed by DMSO and DPBS 4 times, respectively.

3.17. Labeling of intracellular organelles

For ER staining, Alexa Fluor 594-conjugated concanavalin A (ConA) in DPBS (10 μg/mL) was added to the DNFC-stained cells. After 30 min-incubation under the dark and with shaking, the cells were washed with DPBS 4 times. For the staining of mitochondria and NPC, the DNFC-stained cells were incubated with a blocking buffer (3% BSA in DPBS) for 30 min and added with primary antibodies (2 μg/mL) in the blocking buffer. After 3 h, the cells were washed three times with the blocking buffer and further incubated with secondary antibodies (4 μg/mL) in the blocking buffer for 1 h. Finally, the cells were washed with DPBS extensively.

3.18. CLSM observation of the DNFC-stained cells

After staining of fixed cells, CLSM images were acquired by using LSM880 (Carl Zeiss). Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm. Red channel (600-700 nm) images of ConA, TOMM20, and NPC were obtained under excitation at 543 nm, 633 nm and 543 nm, respectively.

3.19. Human tissue sample acquisition

Human tissue samples were collected between February and April in 2018 at the department of plastic and reconstructive surgery in the Boramae Medical Center of Seoul National University. A keloid tissue of the earlobe, a basal cell carcinoma sample of the facial skin, and a dermatofibrosarcoma protuberance sample of the

back were offered for this study. The institutional review board of the Boramae Medical Center reviewed and approved the study protocol (IRB No. 30-2017-11 and 26-2017-20) and all patients provided informed consent agreements. Subject identification was blinded at tissue sampling. Briefly, tissue specimens used surplus tissues remaining after wide excision of skin cancer or keloid resection surgery.

3.20. Production of formalin fixed paraffin-embedded tissues and microsection

On gross, tissues were dissected to about 3 mm in thickness, then tissues were soaked into 4% formalin solution for 1 day. The formalin-fixed tissues were placed into the automatic tissue processor instrument (EFTP) (Intelsint: Villarbasse. Turin, Italy), in which the tissues were submerged in each reagent and permeated through a series of reagents overnight according to the standard protocol.[20] Next day, the tissues were embedded in paraffin wax blocks. Tissue-paraffin blocks (after cooling) were serially microsectioned to 3 μm in the thickness using a rotary microtome apparatus. Each tissue slice was put on a glass slide for the DNFC staining or the conventional histochemical staining.

3.21. H&E- and MT-staining of human tissues for observation with optical microscopy

Hematoxylin and eosin (H&E) stain was performed using the autostainer (Leica autostainer XL) (Leica Biosystems: Newcastle Upon Tyne, UK). The staining procedures were as follows; tissue slices were dewaxed, rehydrated through descending grades of alcohol to water, stain in hematoxylin, washed well in running

tap water, differentiated in acid alcohol, washed well, stained in eosin Y, washed in running tap water, dehydrated through alcohols, and cleared with xylene. As for Masson trichrome (MT) stain, tissues were deparaffinized and rehydrated, then submerged in Boulin's solution overnight. After rinsed with distilled water, tissues were stained by Weigert's iron hematoxylin, washed wall in tap water, treated with phosphomolybdic acid solution, stained with aniline blue solution, washed lightly, treated with acetic acid solution, dehydrated, and cleared in xylene. Finally, each stained-tissue on the glass slide was mounted in permanent mounting medium, and microscopic features were interpreted using an optical microscope.

3.22. Staining of human tissues by the DNFC method

Citric acid monohydrate (0.200 g, 0.952 mmol) was dissolved in DMSO (1.0 mL) and DIPEA (162 μL, 0.952 mmol) was added to the solution. PyBOP (0.496 g, 0.952 mmol) was dissolved in DMSO (2.0 mL) and the solution was added to the citric acid solutions. Human tissue samples in a glass chamber were incubated in the prepared solution for 30 min at room temperature under the dark condition and with shaking. After the staining, the samples were washed with DMSO and DPBS 4 times, respectively.

3.23. CLSM observation of the DNFC-stained tissues

After staining of fixed tissues, CLSM images were acquired by using LSM880 (Carl Zeiss). Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm. Red channel (600-700 nm) images of ConA, TOMM20, and NPC were obtained under excitation at 543 nm, 633 nm and 543 nm, respectively.

4. Results and Discussion

4.1. Facile synthesis of Cit-based fluorophores at mild conditions

Inspired by the proposed mechanism in which TPA formation begins from successive amide/imide couplings between Cit and Cys (Figure 2), we expected that TPA formation could be facilitated by coupling reagents in peptide synthesis (Figure 3). Thus, we checked the effect on fluorophore formation of adding (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP), one of the strongest peptide coupling reagents, to a mixture of Cit and Cys. Furthermore, we analysed the effect of microwave irradiation, which is known to promote amide formation.[21, 22] The detailed methods are given in the supplementary information. Figure 4 compares the fluorescence intensities of Cit/Cys mixtures after treating under different reaction conditions. The mixture heated at 180°C for 1 h using the previous method (control)[14] emitted only minimal fluorescence at 430 nm when excited at 365 nm. Remarkably, when PyBOP was added to a Cit/Cys mixture, the intensity was over 400 times higher after reacting for 1 h at room temperature (25°C). The fluorescence intensity was further increased by microwave irradiation. After 1 h microwave irradiation at 75°C, the PyBOP-containing mixture exhibited approximately 3,500 times higher fluorescence intensity than did the control. The remarkable increase in fluorescence achieved by adding PyBOP confirmed the hypothesis that amide coupling reagents can facilitate the formation of fluorescent structures from Cit and Cys. After 48 h at room temperature, the reaction was nearly saturated (Figure 7C). When it was assisted by both PyBOP and microwave irradiation, increased with increasing temperature (Figure 7A) and approached the maximum value within 30 min at 75°C (Figure 7B).

The major fluorescent reaction product exhibited a ^{1}H NMR spectrum matching that of TPA (Figure 5, Figure 29).[13, 14] The *m/z* value (Figure 36A) and

absorption/emission spectra (Figure 6) also suggested that the fluorescent product of the PyBOP/microwave-assisted reaction is TPA. It was verified that the TPA structure could be synthesised much more efficiently using coupling reagents and microwave irradiation under milder conditions than those previously reported. We also applied the PyBOP/microwave-assisted method to a mixture of L-serine (Ser) and Cit, as Ser, the hydroxyl analogue of Cys, reportedly emitted long wavelength fluorescence after the reaction with Cit.[12] The resulting fluorescence intensity was about 10 times higher than that of the sample prepared by heating at 180°C (Figure 8A). The fluorescence of Ser-based 5-oxo-2,3-dihydro-*5H*-[1,3]oxazolo[3,2-*a*]pyridine-3,7-dicarboxylic acid (OPA) was red-shifted by about 40–50 nm compared to that of Cys-based TPA (Figure 8B).

4.2. Fluorophore formation on *N*-termini of peptides

We then determined whether the TPA fluorophore could be *de novo* generated on a peptide structure. Since both amine and thiol groups are needed for dihydrothiazolopyridone ring formation, *N*-terminal Cys with both groups in a peptide sequence may be a suitable target for TPA introduction. A model dipeptide, cysteine-leucine-*n*-butylamine (Cys-Leu-Bu), was reacted with Cit at room temperature in the presence of PyBOP. The ^{1}H NMR (Figure 10, Figure 35) and LC/MS spectra (Figure 36B) confirmed that the TPA structure was successfully introduced to the peptide *N*-terminus. Like TPA, the TPA-conjugated dipeptide (compound **7**) emitted strong fluorescence at 420 nm when excited at 350 nm (Figure 11). The TPA fluorophore was also generated on longer peptides conjugated to a solid phase peptide synthesis resin.

We synthesised a 12-meric TAT peptide (CYGRKKRRQRRR) and a 17-meric peptide (LK-peptide) (CLKKLCKLLKKLCKLAG), both with *N*-terminal cysteines,

on a solid phase resin via typical 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry (Figure 12). The peptides were treated with Cit/PyBOP and cleaved from the resin. Two fluorescent peptides were purified by HPLC and showed *m/z* values close to those calculated for TPA-containing TAT and LK-peptides (Figure 13, Figure 37). TPA-TAT and TPA-LK peptides exhibited maximum emission intensities at 440 and 423 nm, respectively (Figure 14A), suggesting that the emission is dependent upon the peptide sequences. The hydrophobic leucine (L) residues next to the cysteine (C) in the LK peptide may induce the blue shift of the emission maximum from that of free TPA (430 nm) through hydrophobic interaction,[23] whereas the tyrosine (Y) and successive hydrophilic charged residues in the TAT sequence may induce the red shift.[24] Moreover, the fluorescence intensity of the fluorophore was strongly affected by the peptide bond between the carboxylate on the thiazolidine ring and the next amino acid, as confirmed by the 30% decrease in the fluorescence intensity of compound **7** (Figure 14B) with respect to free TPA.

A peptide having an *N*-terminal TPA structure could be directly used for bioimaging. As the TAT and LK peptides were reported as cell-penetrating peptides, [24, 25] we observed the intracellular distribution of blue fluorescence in *HeLa* cells treated with the TPA-peptides. Confocal laser scanning microscopy (CLSM) images show blue fluorescent dots in the cytoplasm on two-photon excitation at 720 or 780 nm, according to the maximum excitation wavelengths of TPA-LK and TPA-TAT, respectively (Figure 15). The TPA-TAT-treated cells showed concentrated localisation of the TPA fluorescence in the nucleus (Figure 16).[26] These results suggested that the cell-penetrating activity of both peptide sequences as well as the nuclear localisation behaviour of the TAT sequence was well maintained even after TPA formation on the peptides.

4.3. Fluorophore generation in cultured cells

We expected that proteins in biological samples could be directly labelled with the fluorophore via the same method using Cit and PyBOP. As a proof of concept, we treated fixed *HeLa* cells with Cit and PyBOP in a mixed solvent (DMSO/H_2O = 70/30) (Figure 17). As shown in CLSM images, the cells exhibit both blue and green fluorescence on illumination with a 405 nm laser (Figure 18). After 30 min incubation with Cit and PyBOP, fluorescent dots appeared in the cytoplasm of the cells, and the fluorescence intensity gradually increased with the incubation time. Clearly, the fluorescence-generation efficiency depended on the staining solvent. Without organic solvent, fluorescence was negligible, probably due to inefficient dehydration against fluorophore formation together with poor water-solubility of PyBOP. However, the fluorescence developed much faster in organic solvents including *N*,*N*-dimethylformamide (DMF) or dimethyl sulfoxide (DMSO) (Figure 19A). Furthermore, while blue and green puncta were discretely distributed mainly in the cytoplasm when the cells were stained in DMSO/H_2O, the fluorescence was distributed throughout the cell interior under staining in organic solvents (Figure 20A). Fluorophore formation may be better promoted in anhydrous environments, such as membrane structures in aqueous solvents, to show discrete fluorescent puncta. However, we suggested that fluorophores may be indiscriminately and rapidly generated on proteins, even in non-membrane structures, in organic solvents.

We performed a more detailed analysis of the intracellular fluorescence distribution from *HeLa* cells stained by the DNFC method using DMSO/H_2O. Interestingly, at the focal plane near the nucleus in the magnified CLSM images (Figure 21), both blue and green fluorescent dots were mainly observed in the cytoplasm outside the nucleus.[27] Moreover, we compared the co-localisation of DNFC-derived blue and green fluorescence signals with red fluorescence-labelled

molecules targeting specific membrane proteins in major membrane-bound organelles: the endoplasmic reticulum (ER), mitochondria and nuclei. Co-localisation was identified as purple or yellow signals in the merged images (Figure 22), clearly indicating a high extent of co-localisation of DNFC-derived green fluorescence and ER-targeting concanavalin A (Con A)-labelled red fluorescence. Statistical image analysis of over 500 points in cells also confirmed a considerably higher degree of fluorescence co-localisation of DNFC-derived and ER-targeting molecules than those targeting mitochondrial and nuclear membranes (Figure 23).

The higher degree of DNFC staining on the ER may be rationalised by the features of TPA or OPA formation. Considering that an amino group is needed as well as the thiol or hydroxyl group for formation of the thiazolidine or oxazolidine ring of TPA and OPA, we predicted that *N*-terminal Cys or Ser would be the optimised point for generating the fluorophores under the mild reaction conditions of DNFC staining. Protein translation is initiated from L-methionine (Met) by ribosomes; thus, most cytosolic proteins initially have an *N*-terminal Met. Post-translational cleavage may subsequently occur to expose a non-Met residue on the *N*-termini of certain proteins. Cleavage of *N*-terminal signal peptides in the ER, which directs intracellular protein trafficking, is a well-known post-translational cleavage reaction. Most non-cytosolic proteins such as membrane, nuclear, mitochondrial and secreted proteins are trimmed by endopeptidases in the ER membrane and sorted to their final destinations through the ER and Golgi complex.[28-30] Thus, compared to cytosolic proteins, ER proteins are more likely to expose non-Met amino acids, conceivably Cys or Ser, on the *N*-termini, where DNFC staining may occur at higher rates. A portion of Cys- or Ser-terminal proteins in membrane-bound organelles, which were predominantly trafficked from the ER, may also react with Cit to display DNFC-based fluorescence in the organelles (Figure 22). The CLSM images at the focal plane near the plasma membrane also

showed strong fluorescence on the fibrous structures outside the cells, presumably extracellular matrix proteins, which were also secreted from the ER (Figure 20B).

Remarkably, green fluorescence is much stronger than blue at an excitation wavelength of 405 nm in all DNFC-based cell staining images. The stronger green fluorescence was still observed in the CLSM images on two-photon excitation at 780 nm (Figure 20C). When we incubated the same amount of Cys or Ser with Cit and PyBOP in DMSO/H_2O for 30 min, the Cys-containing mixture emitted about 5.5 times stronger fluorescence than the Ser-containing mixture at their maximum excitation wavelengths (Figure 19B). However, when excited at 405 nm, the Cys-containing mixture showed approximately 1.7 times higher blue fluorescence intensity (450 nm), whereas the Ser-containing mixture showed 1.5 times higher green fluorescence intensity (550 nm) (Figure 19C). When we incubated Met in the cell-free DNFC staining condition, the Met-containing mixture emitted negligible blue and green fluorescence by the excitation at 405 nm. In comparing the images and spectra, we inferred that the green fluorescence mainly originated from Ser-based OPA and blue from Cys-based TPA in the DNFC-stained cells. Additionally, we anticipated that the predominance of green fluorescence in the cell staining was due to the different frequencies of Cys- or Ser-terminal proteins in cells. The whole *N*-terminal proteome of human proteins has not yet been identified; however, Ser is frequently exposed in a non-acetylated form at the *N*-termini of trimmed proteins after cleavage of signal peptides. Furthermore, the exposure of *N*-terminal Cys is extremely rare.[29, 31] Variation in the green/blue fluorescence ratios might be linked to variation in Cys- and Ser-terminal proteomes among organelles. We propose that the *N*-terminal proteome in mitochondria might have a higher Cys-/Ser- ratio than the ER or nuclear membrane (Figure 23).

We compared the morphologies of *HEK293T* (human embryonic kidney),[32-34] *THP-1M* (human macrophage)[35, 36] and *NIH3T3* (mouse fibroblast)[37, 38] as other

representative cell lines after staining with the DNFC method (Figure 24). The CLSM image shows the characteristic morphology of each cell line: the extended multipolar *HEK293T* and *NIH3T3* and irregular spherical *THP-1M* cells. Both blue and green fluorescence preferentially developed in the cytoplasm rather than in the nuclei of the three cell lines, as with *HeLa* cells. *HEK293T* cells developed the lowest fluorescence intensities at both green and blue wavelengths. However, both *THP-1M* and *NIH3T3* cells showed strong fluorescence with DNFC staining. Particularly, *THP-1M* cells exhibited much stronger blue fluorescence than other cells. Both the Cys- and Ser-terminal proteomes and the protein expression levels in ER might be characteristically represented by variations in the fluorescence intensity and green/blue fluorescence in each cell line.

4.4. DNFC staining of human tissues

Inspired by each cell type exhibiting a characteristic blue and green fluorescence distribution after DNFC staining, we predicted that detailed shapes and structures of human tissues could be identified by this staining, as with conventional histochemistry. Human skin and attached subcutaneous fat tissues were fixed with 4% formalin solution and embedded in paraffin blocks. Tissue slices of thickness 3 μm were microsectioned from the paraffin blocks, placed on glass slides and treated with Cit and PyBOP in DMSO, rather than DMSO/H_2O, to enhance reagent penetration into the densely packed epidermis.[39] DNFC-stained skin/subcutaneous tissue and basal cell carcinoma (BCC) exhibited strong blue and green emission when excited at 405 nm, which was much stronger than the autofluorescence of tissues (Figure 25, 26).[40] In the epidermis, multiple layers of flake-like keratin at the stratum corneum, the densely packed uppermost layer, were clearly observed (Figure 27A, H2).[41, 42] Additionally, we confirmed that the extracellular matrix

surrounding polyhedral keratinocytes in the stratum spinosum (Figure 27A, H1), [41, 43] the lower epidermal layer, was strongly stained by the DNFC method. As the paraffin-tissue blocks were serially microsectioned into slices, the DNFC-stained tissue slice and adjacent tissue slices of conventional histochemical stains, hematoxylin and eosin stain (H&E stain) and Masson's trichrome stain (MT stain), exhibited almost the same microscopic features of substructures, allowing the DNFC-stained substructures to be histomorphologically interpreted (Figure 27B). In the dermis (Figure 27B, top panel), the DNFC could stain collagen and elastic fibre bundles.[44] In the subcutaneous fat tissues (Figure 27B, bottom panel), the extracellular matrix and thin cytoplasm, except for the lipid droplets of adipocytes, emitted fluorescence with DNFC staining.[45] The thin cytoplasm of adipocytes was visually accentuated in the DNFC-stained sample over the H&E- or MT-samples.

We also discerned substructures of a DNFC-stained skin/subcutaneous tissue sample from a patient with BCC. Mesh-like cribriform structures were observed in a DNFC-stained image of sebaceous glands (Figure 28A, top panel).[46] As with the adipocytes, lipid-based adipocyte material was not stained by the DNFC method. Other images showed the characteristic shape of eccrine glands (Figure 28A, middle panel).[47] Interestingly, we clearly observed bright dot-like structures scattered throughout the DNFC-stained eccrine glands, although not in the H&E- or MT-stained samples. The bright fluorescence in the glands indicates that they have large quantities of proteins with *N*-terminal Ser or Cys, and the dot-like structure may be a characteristic feature of the eccrine gland. Notably, tumour cells were strongly stained by the DNFC method (Figure 28A, bottom panel). The array of tumour cells and unstained stroma are characteristic histomorphological features of BCC, which was also verified by the H&E- and MT-stained images.[48] The ratio of green and blue fluorescence intensities was strongly dependent on the substructure of the tissue, suggesting variation in the *N*-terminal proteomes. In another skin/connective tissue

sample, an arteriole with a thicker wall and a narrower lumen than capillaries was identified (Figure 28B, left panel).[49] In another point in the sample, fat necrosis was distinguished by a structure of unstained fat enclosed by histiocytes and multinucleated giant cells (Figure 28B, right panel).[50]

We showed that a simple 30 min treatment with Cit and PyBOP is sufficient for staining specific sections of human tissue samples by forming green and blue fluorophores, probably depending on the protein distribution. We can examine arrays of cells and extracellular matrices of the DNFC-stained sample at sufficiently high resolutions to determine the substructure and identify abnormal tissues, as with conventional H&E and MT staining-based histology. Although we need to identify the specific substructures stained by the DNFC method in more detail and analyse the characteristic proteomes in the substructure in future research, we believe that the DNFC-based analysis could be a strong alternative diagnosis method for supporting conventional histology by providing additional information concerning the protein distribution.

5. Conclusions

We report that strong blue and green fluorophores based on a fused pyridone ring structure are *de novo* generated on peptides and protein *N*-termini from non-fluorogenic Cit and amino acids, Cys or Ser, under mild conditions. An amide coupling reagent, PyBOP, and microwave irradiation strongly facilitate fluorophore formation even at low temperatures. The fluorophores can be generated on the *N*-terminal Cys of peptides synthesised in either the liquid or solid phase. Moreover, intracellular or extracellular proteins with *N*-terminal Cys or Ser are fluorescently labelled by the DNFC method. ER is more strongly labelled than other intracellular organelles by DNFC staining, probably because it is rich in proteins with *N*-terminal Ser. The DNFC method, which exhibits characteristic distributions of fluorescently labelled proteins in biosamples, could be applied to human tissue staining. Rapid staining of paraffinized human tissues with Cit and PyBOP allowed substructure identification and evaluation of tissue abnormalities. The DNFC-staining characteristics of cells and tissues in this study are summarized in Table 1 and 2 for such future applications. As a new strategy for introducing fluorescence to biological molecules, the DNFC method is expected to contribute to bioimaging and diagnosis. Fluorophore formation from non-fluorogenic biomolecules may aid discovery of new chemical or biological fluorophores.

6. References

[1] B. N. Giepmans, S. R. Adams, M. H. Ellisman and R. Y. Tsien, *Science*, 2006, **312**, 217-224.

[2] M. Fernandez-Suarez and A. Y. Ting, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 929-943.

[3] T. D. Ashton, K. A. Jolliffe and F. M. Pfeffer, *Chem. Soc. Rev.*, 2015, **44**, 4547-4595.

[4] M. S. Goncalves, *Chem. Rev.*, 2009, **109**, 190-212.

[5] I. L. Medintz, H. T. Uyeda, E. R. Goldman and H. Mattoussi, *Nat. Mater.*, 2005, **4**, 435-446.

[6] M. Han, X. Gao, J. Z. Su and S. Nie, *Nat. Biotechnol.*, 2001, **19**, 631-635.

[7] L. Sun, T. Liu, H. Li, L. Yang, L. Meng, Q. Lu and J. Long, *ACS Appl. Mater. Interfaces*, 2015, **7**, 4990-4997.

[8] L. Gao, J. Q. Wang, L. Huang, X. X. Fan, J. H. Zhu, Y. Wang and Z. G. Zou, *Inorg. Chem.*, 2007, **46**, 10287-10293.

[9] S. Kim, T. Y. Ohulchanskyy, H. E. Pudavar, R. K. Pandey and P. N. Prasad, *J. Am. Chem. Soc.*, 2007, **129**, 2669-2675.

[10] R. N. Day and M. W. Davidson, *Chem. Soc. Rev.*, 2009, **38**, 2887-2921.

[11] M. Chalfie, Y. Tu, G. Euskirchen, W. Ward and D. Prasher, *Science*, 1994, **263**, 802-805.

[12] J. Yang, Y. Zhang, S. Gautam, L. Liu, J. Dey, W. Chen, R. P. Mason, C. A. Serrano, K. A. Schug and L. Tang, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 10086-10091.

[13] W. Kasprzyk, S. Bednarz and D. Bogdal, *Chem. Commun.*, 2013, **49**, 6445-6447.

[14] W. Kasprzyk, S. Bednarz, P. Żmudzki, M. Galica and D. Bogdał, *RSC Advances*, 2015, **5**, 34795-34799.

[15] Z. Xie, J. P. Kim, Q. Cai, Y. Zhang, J. Guo, R. S. Dhami, L. Li, B. Kong, Y. Su, K. A. Schug and J. Yang, *Acta. Biomater.*, 2017, **50**, 361-369.

[16] M. J. Krysmann, A. Kelarakis, P. Dallas and E. P. Giannelis, *J. Am. Chem. Soc.*, 2012, **134**, 747-750.

[17] Y. Song, S. Zhu, S. Zhang, Y. Fu, L. Wang, X. Zhao and B. Yang, *Journal of Materials Chemistry C*, 2015, **3**, 5976-5984.

[18] S. L. Pedersen, A. P. Tofteng, L. Malik and K. J. Jensen, *Chem. Soc. Rev.*, 2012, **41**, 1826-1844.

[19] G. S. Vanier, *Methods Mol. Biol.*, 2013, **1047**, 235-249.
[20] L. T. Specer, J. D. Bancroft, W. G. Jones, *In Bancroft's theory and practice of histological techniques*, Elsevier, Churchill Livingstone, UK, 2013.
[21] L. Perreux, A. Loupy and F. Volatron, *Tetrahedron*, 2002, **58**, 2155-2162.
[22] R. R. Poondra and N. J. Turner, *Org. Lett.*, 2005, **7**, 863-866.
[23] X. Yang, R. Lu, H. Zhou, P. Xue, F. Wang, P. Chen and Y. Zhao, *J. Colloid Interface Sci.*, 2009, **339**, 527-532.
[24] F. Milletti, *Drug Discov. Today*, 2012, **17**, 850-860.
[25] S. Jang, S. Hyun, S. Kim, S. Lee, I. S. Lee, M. Baba, Y. Lee and J. Yu, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 10086-10089.
[26] L. Pan, Q. He, J. Liu, Y. Chen, M. Ma, L. Zhang and J. Shi, *J. Am. Chem. Soc.*, 2012, **134**, 5722-5725.
[27] S. Dewitt, R. L. Darley and M. B. Hallett, *J. Cell Biol.*, 2009, **184**, 197-203.
[28] S. F. Nothwehr and J. I. Gordon, *Bioessays*, 1990, **12**, 479-484.
[29] R. J. Folz, S. F. Nothwehr and J. I. Gordon, *J. Biol. Chem.*, 1988, **263**, 2070-2078.
[30] L. Ellgaard and A. Helenius, *Nat. Rev. Mol. Cell Biol.*, 2003, **4**, 181-191.
[31] A. Martinez, J. A. Traverso, B. Valot, M. Ferro, C. Espagne, G. Ephritikhine, M. Zivy, C. Giglione and T. Meinnel, *Proteomics*, 2008, **8**, 2809-2831.
[32] C. W. Tan, B. S. Gardiner, Y. Hirokawa, D. W. Smith and A. W. Burgess, *BMC Syst. Biol.*, 2014, **8**, 44.
[33] R. Sato, J. Kozuka, M. Ueda, R. Mishima, Y. Kumagai, A. Yoshimura, M. Minoshima, S. Mizukami and K. Kikuchi, *J. Am. Chem. Soc.*, 2017, **139**, 17397-17404.
[34] H. J. Kweon, D. I. Kim, Y. Bae, J. Y. Park and B. C. Suh, *Sci. Rep.*, 2016, **6**, 30684.
[35] S. De Clercq, C. Boucherie, J. Vandekerckhove, J. Gettemans and A. Guillabert, *PLoS One*, 2013, **8**, e78108.
[36] O. Lunov, T. Syrovets, C. Loos, J. Beil, M. Delacher, K. Tron, G. U. Nienhaus, A. Musyanovych, V. Mailander, K. Landfester and T. Simmet, *ACS Nano*, 2011, **5**, 1657-1669.
[37] X. Zhang, L. Dai, A. Wang, C. Wolk, B. Dobner, G. Brezesinski, Y. Tang, X. Wang and J. Li, *Sci. Rep.*, 2015, **5**, 16559.
[38] Y. Morioka, J. Monypenny, T. Matsuzaki, S. Shi, D. B. Alexander, H. Kitayama

and M. Noda, *Oncogene*, 2009, **28**, 1454-1464.
[39] T. Kurihara-Bergstrom, G. L. Flynn and W. I. Higuchi, *Journal of Investigative Dermatology*, 1987, **89**, 274-280.
[40] M. Monici, *Biotechnol. Annu. Rev.*, 2005, **11**, 227-256.
[41] S. Yang, Y. Sun, Z. Geng, K. Ma, X. Sun and X. Fu, *Int. J. Mol. Med.*, 2016, **37**, 1263-1273.
[42] V. Ahlgrimm-Siess, M. Laimer, H. S. Rabinovitz, M. Oliviero, R. Hofmann-Wellenhof, A. A. Marghoob and A. Scope, *Curr. Dermatol. Rep.*, 2018, **7**, 105-118.
[43] T. Sutterlin, E. Tsingos, J. Bensaci, G. N. Stamatas and N. Grabe, *Sci. Rep.*, 2017, **7**, 43472.
[44] G. M. Bove, M. Y. Harris, H. Zhao and M. F. Barbe, *J. Neurol. Sci.*, 2016, **361**, 168-180.
[45] S. Sindhu, R. Thomas, P. Shihab, D. Sriraman, K. Behbehani and R. Ahmad, *PLoS One*, 2015, **10**, e0133494.
[46] D. Frances, N. Sharma, R. Pofahl, M. Maneck, K. Behrendt, K. Reuter, T. Krieg, C. A. Klein, I. Haase and C. Niemann, *Oncogene*, 2015, **34**, 5505-5512.
[47] Y. Gao, M. Li, X. Zhang, T. Bai, G. Chi, J. Y. Liu and Y. Li, *Int. J. Mol. Med.*, 2014, **34**, 997-1003.
[48] P. Saldanha, P. R. Shanthala and K. Upadhaya, *Archives of Medicine and Health Sciences*, 2015, **3**.
[49] B. S. Soh, S. Y. Ng, H. Wu, K. Buac, J. H. Park, X. Lian, J. Xu, K. S. Foo, U. Felldin, X. He, M. Nichane, H. Yang, L. Bu, R. A. Li, B. Lim and K. R. Chien, *Nat. Commun.*, 2016, **7**, 10774.
[50] K. Karochristou, T. Siahanidou, T. Kakourou-Tsivitanidou, K. Stefanaki and H. Mandyla, *J. Perinatol.*, 2006, **26**, 64-66.

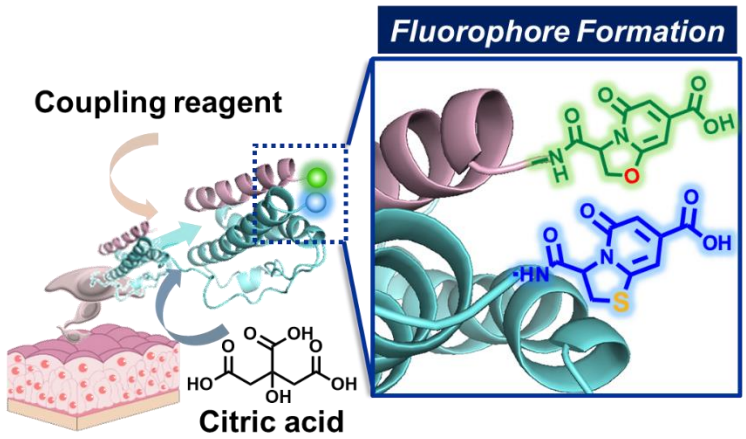


Figure 1. A scheme for formation of pyridone-based TPA and OPA structure by the reaction between Cit and Cys or Ser on proteins in cells and tissues.

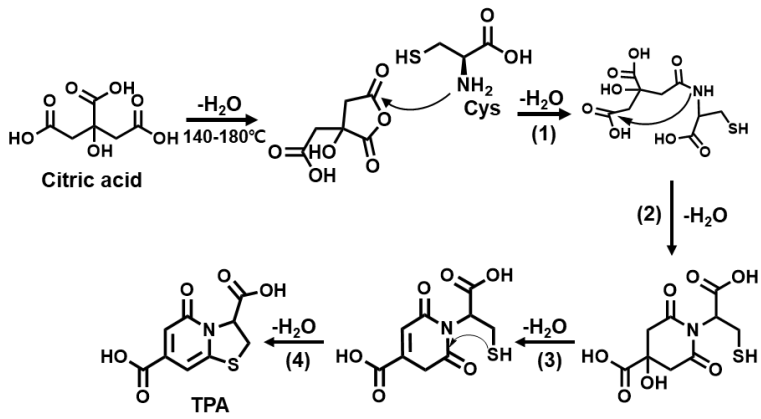


Figure 2. A proposed mechanism of TPA formation without coupling reagents. (1) amide formation, (2) imide formation, (3) dehydration, (4) intramolecular condensation.

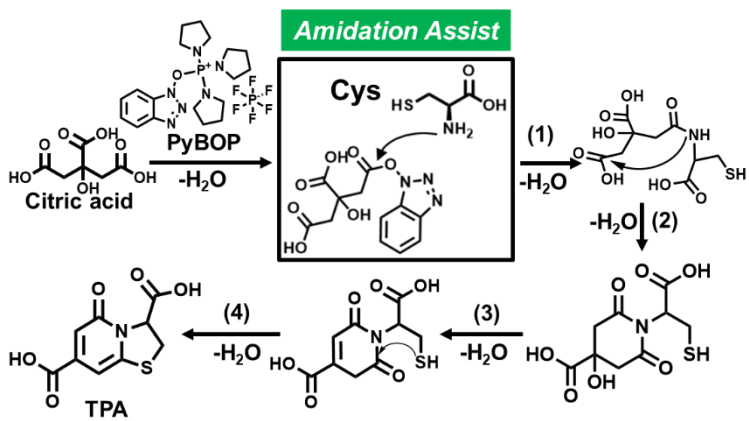


Figure 3. A proposed mechanism of TPA formation with PyBOP. (1): amide formation, (2): imide formation, (3): dehydration, (4): intramolecular condensation.

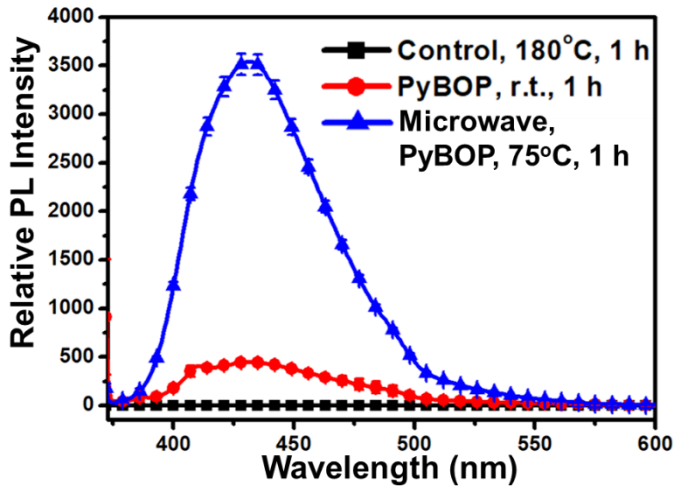


Figure 4. Emission spectra (λ_{ex} = 365 nm) of mixtures containing the same amount of Cit and Cys after various treatments for 1 h: heating at 180°C (control); incubation with PyBOP at 25°C; incubation with PyBOP in presence of microwave irradiation. Each data point represents average ± S.D. (n = 4).

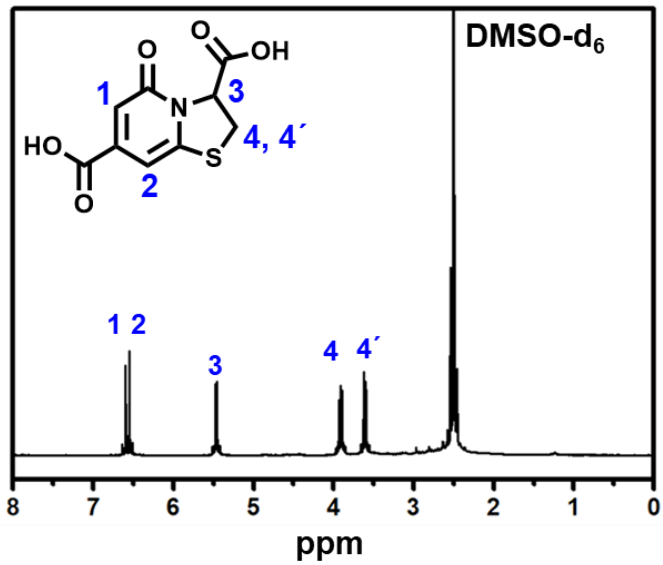


Figure 5. ^{1}H NMR spectrum (DMSO-d_6, 500 MHz) of TPA prepared by the PyBOP-assisted synthesis: δ 6.64-6.60 (s, 1H), 6.55-6.51 (s, 1H), 5.51-5.42 (d, 1H), 3.97-3.85 (m, 1H), 3.66-3.56 (m, 1H).

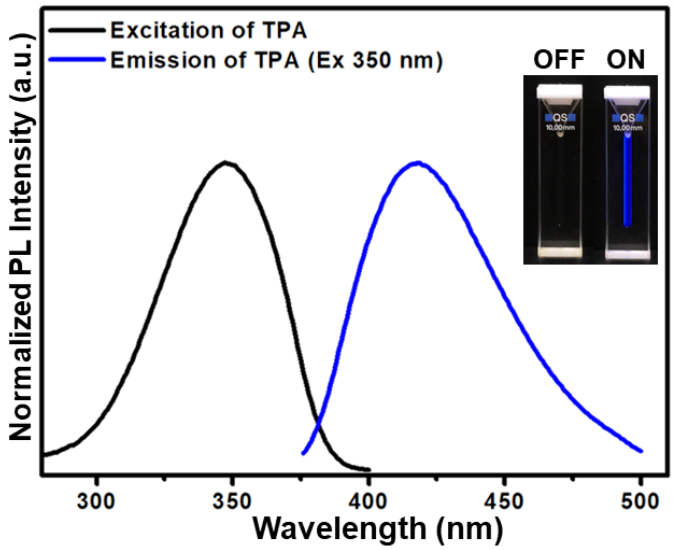


Figure 6. Absorption and emission spectra of TPA prepared by the PyBOP-assisted synthesis in deionized water at 25ºC. (solvent: deionized water, concentration: 0.2 mg/mL). For the emission spectrum, the excitation wavelength was 350 nm. Inset: Fluorescence of TPA illuminated by UV light (365 nm) (solvent: deionized water, concentration: 0.2 mg/mL).

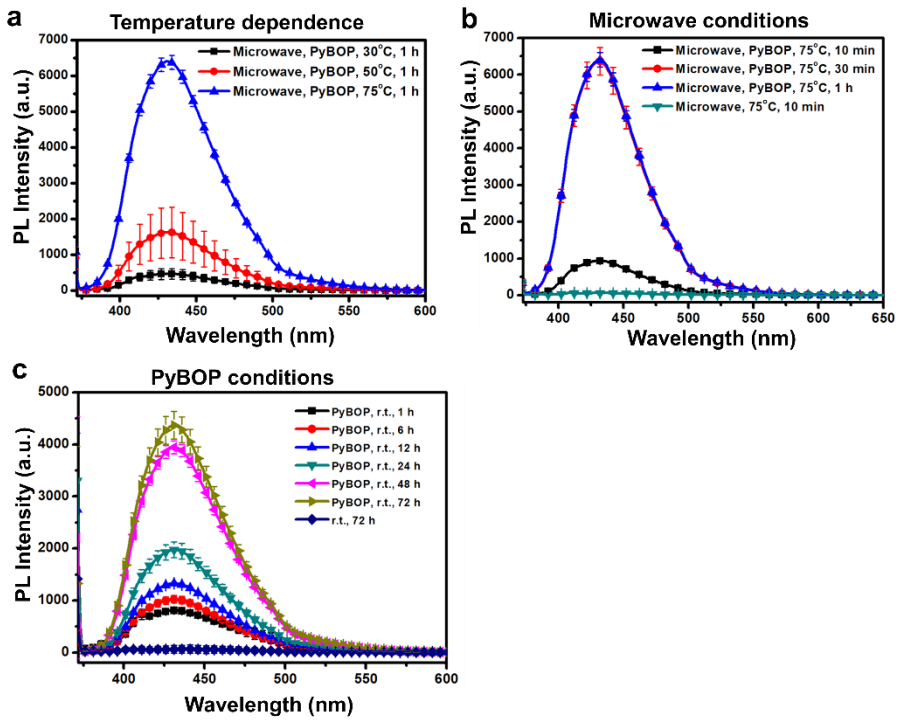


Figure 7. Emission spectra (λ_{ex}= 365 nm) of mixtures in DMF containing the same amount of citric acid and L-cysteine which were treated with various reaction conditions: (a) with microwave irradiation at various reaction temperatures (black: 30ºC, red: 50ºC, blue: 75ºC); (b) with microwave irradiation for various reaction time at 75°C (black: 10 min, red: 30 min, blue: 1 h, cyan: 10 min without PyBOP); (c) without microwave irradiation for various reaction time at 25°C. (black: 1 h, red: 6 h, blue: 12 h, cyan: 24 h, magenta: 48 h, brown: 72 h, navy: 72 h without PyBOP). (a-c) Each data point represents the average ± S.D. (n = 4).

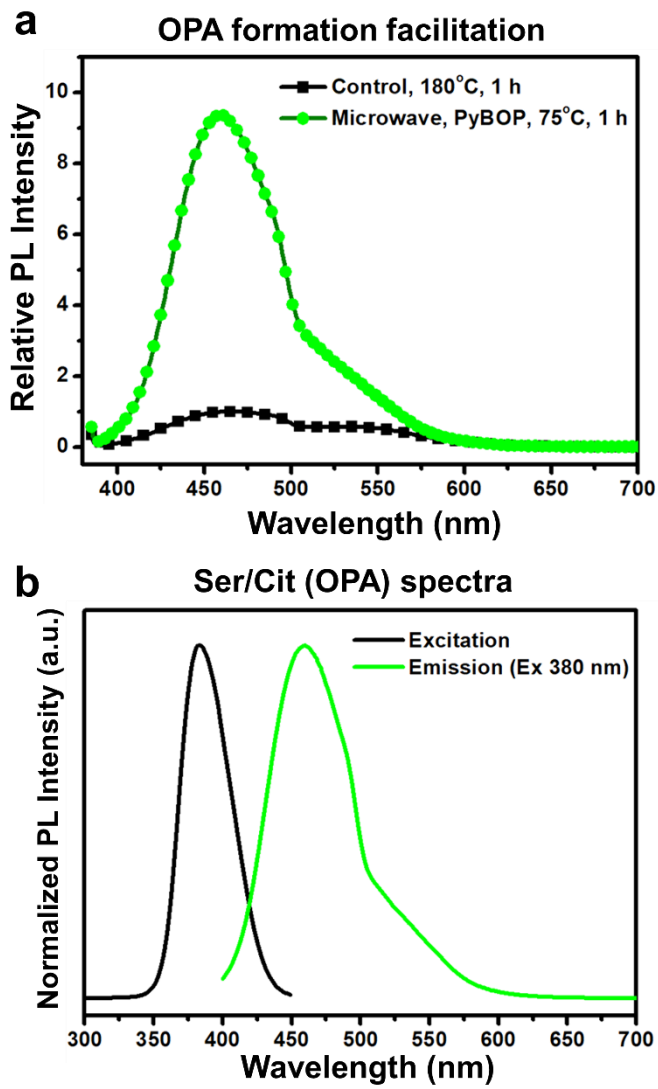


Figure 8. (a) Emission spectra (λ_{ex} = 380 nm) of a mixture of citrate (0.476 mmol) and L-serine (0.48 mmol) in DMF (2.0 mL) for 1 h at 180°C (control) or incubation with PyBOP (0.476 mmol) and DIPEA (0.476mmol) in presence of microwave irradiation for 1 h at 75°C. (b) Absorption and emission spectra (λ_{ex} = 380 nm) of a mixtures containing the same amount of citrate and L-serine in DMF after treatment in the same condition of (a).

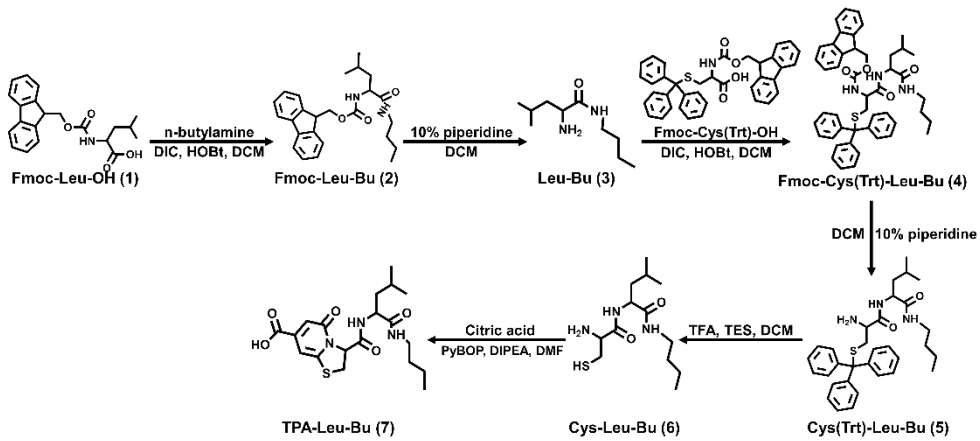


Figure 9. Synthesis of TPA-Leu-Bu **(7)**; model peptide in liquid phase.

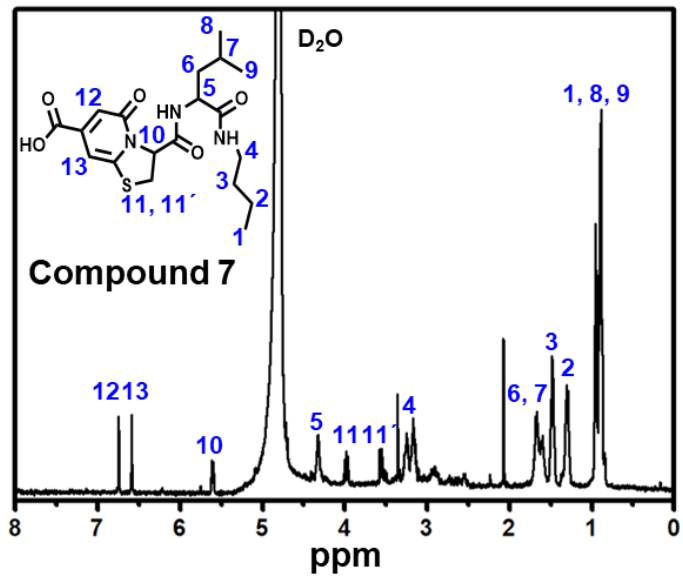


Figure 10. ^{1}H NMR spectrum (D_2O, 500 MHz) of Cys-Leu-Bu with a TPA structure (compound **7**): δ 6.75 (s, 1H), 6.59 (s, 1H), 5.61-5.60 (d, 1H), 4.33-4.32 (m, 1H), 4.00-3.96 (m, 1H), 3.57-3.55 (m, 1H), 3.25-3.15 (m, 2H), 1.68-1.29 (m, 7H), 0.96-0.88 (m, 9H), 2.07 (acetonitrile), 3.36 (methanol).

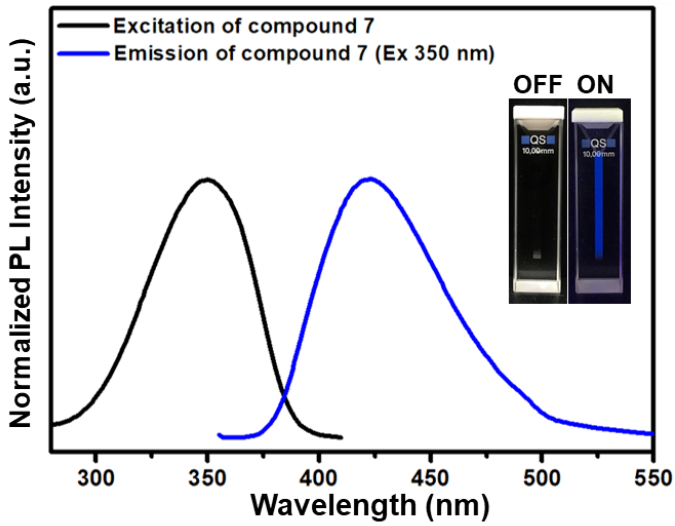


Figure 11. Absorption and emission spectra of compound **7** (solvent: deionized water, concentration: 0.3 mg/mL) in deionized water. Inset: Fluorescence of compound **7** illuminated by UV light (365 nm) (solvent: deionized water, concentration: 0.3 mg/mL).

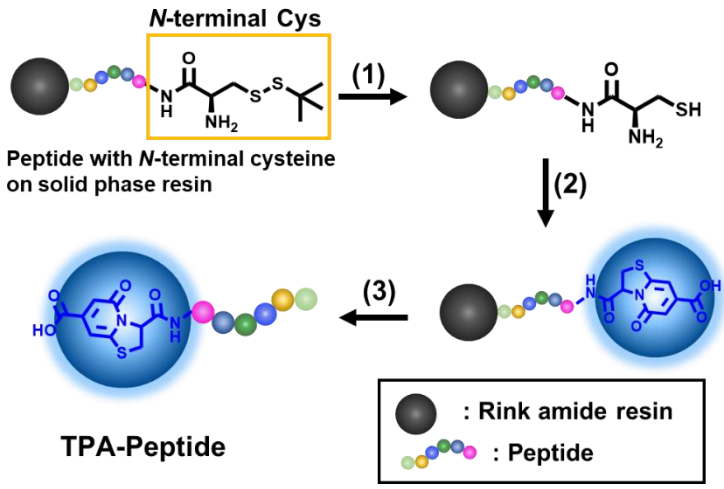


Figure 12. Synthetic scheme of TPA-peptides on SPPS resin. (1): deprotection of S-(*tert*-butylthio) group, (2): formation of TPA on the peptide, (3): cleavage of the TPA-peptide from the resin.

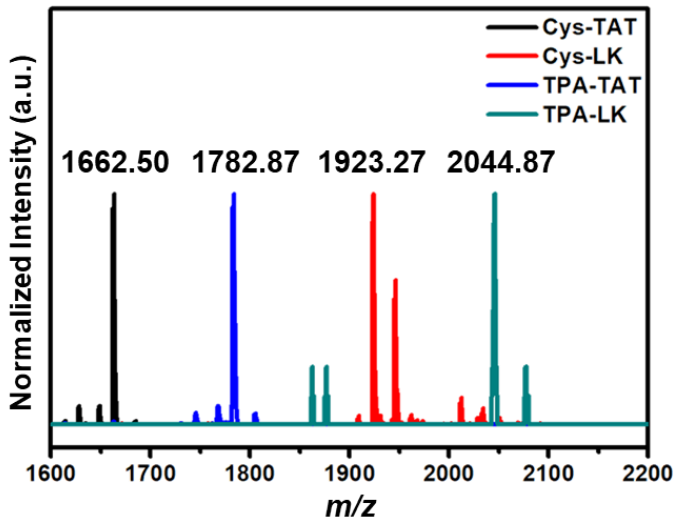


Figure 13. MALDI-TOF spectra of Cys-TAT, Cys-LK, TPA-TAT and TPA-LK peptides. MS (*m/z*): Cys-TAT (calculated: 1782.94; found: 1782.87), Cys-LK (calculated: 1923.35; found: 1923.27), TPA-TAT (calculated: 1782.94; found: 1782.87) and TPA-LK (calculated: 2043.30; found: 2044.87).

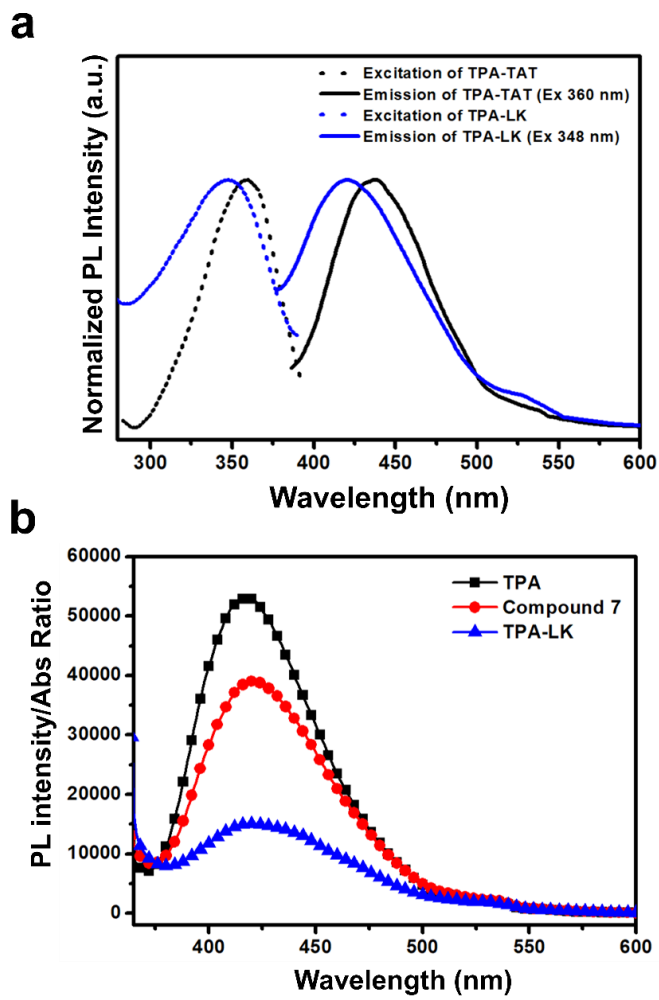


Figure 14. (a) Absorption and emission spectra of TPA-TAT and TPA-LK peptides in deionized water. (b) PL intensity/Absorbance ratios of TPA, compound **7** and TPA-LK peptide in deionized water. Each sample concentration was 20 μM. Absorbance was measured at 350 nm and the excitation wavelength was 350 nm.

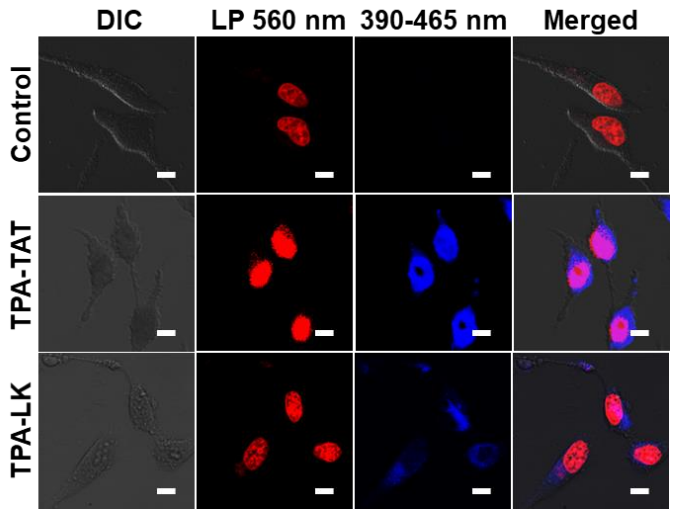


Figure 15. CLSM images of *HeLa* cells treated with TPA-TAT (5 μM) and TPA-LK (500 nM) for 12 h. Cell nuclei were counter-stained with a red fluorescent SYTO 59 dye. Red channel (560 nm long-pass filter) images were obtained by excitation with a 1 mW laser at 543 nm and blue channel (390-465 nm) images were done by two-photon excitation with 2 mW lasers at 780 nm (TPA-TAT) or 720 nm (TPA-LK), respectively. All scale bars represent 10 μm.

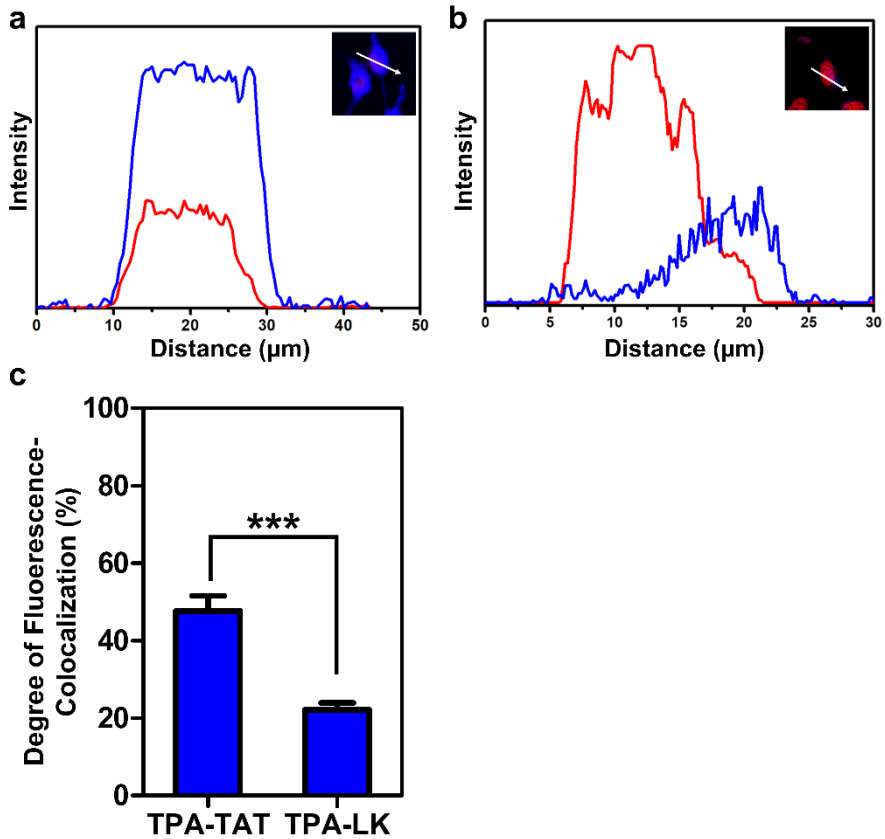


Figure 16. Localization of TPA-labeled peptides in *HeLa* cells. Line profiles indicating the fluorescence distribution on *HeLa* cells treated with (a) TPA-TAT and (b) TPA-LK. The blue lines indicate the intensity profiles of blue fluorescence (390-465 nm) from TPA and the red lines indicate those of red fluorescence (560 nm) from SYTO 59 through the white arrow lines in the CLSM images of cells. (c) Comparison of the colocalization ratio of blue fluorescence with red fluorescence in *HeLa* cells. The colocalization ratio was calculated from over 500 points in a CLSM image of *HeLa* cells by the Imaris microscopy analysis software. Total four images (over 2,000 points) were analyzed and the values were averaged. Each bar represents the average ± S.D. ($n = 4$). The *** mark means $p < 0.0001$ through the Student's *t*-test.

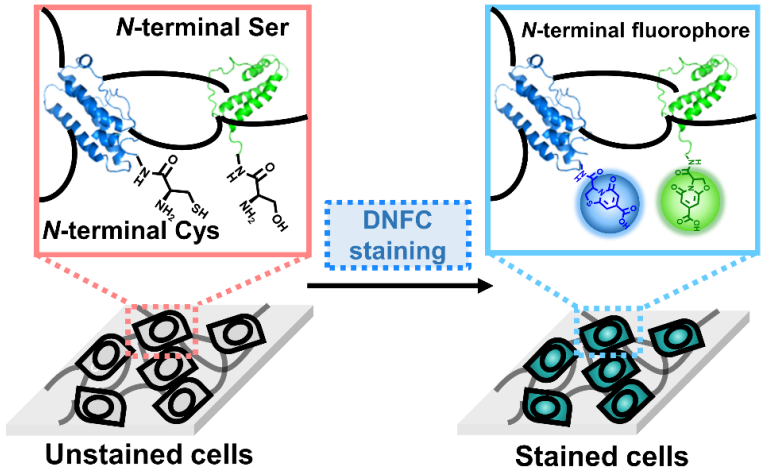


Figure 17. Simplified scheme of DNFC staining on cells.

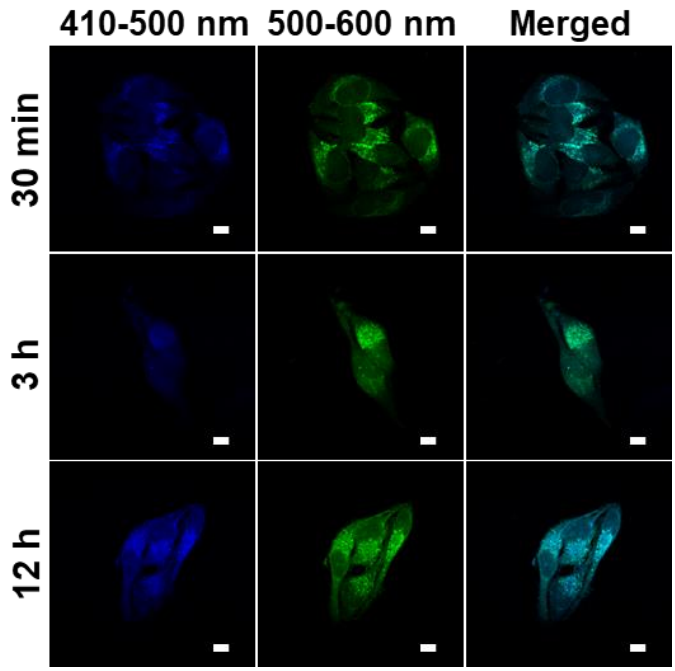


Figure 18. Time-dependent development of fluorescence in *HeLa* cells stained by the DNFC method. Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm with a laser power of 5 mW. All scale bars represent 10 μm.

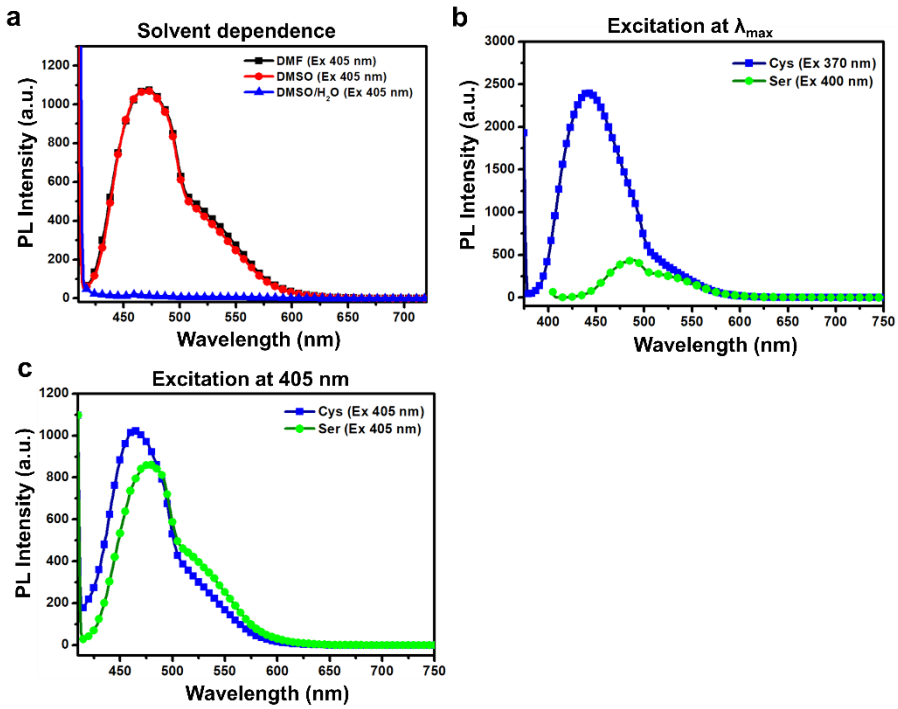


Figure 19. (a) Emission spectra (λ_{ex}= 405 nm) of mixtures containing the same amount of citric acid and L-cysteine after incubation with PyBOP for 30 min at 25°C in various reaction solvents (black: DMF, red: DMSO, blue: DMSO/H_2O). (b) Comparison of fluorescence emission by excitation at the maximum absorption wavelength (maximum λ_{ex} of Cys/Cit= 370 nm and Ser/Cit= 400 nm) from a Cys/Cit mixture and a Ser/Cit mixture after incubation with PyBOP for 30 min at 25°C in DMSO/H_2O. (c) Comparison of fluorescence emission by excitation at 405 nm from a Cys/Cit mixture and a Ser/Cit mixture after incubation with PyBOP for 30 min at 25°C in DMSO/H_2O.

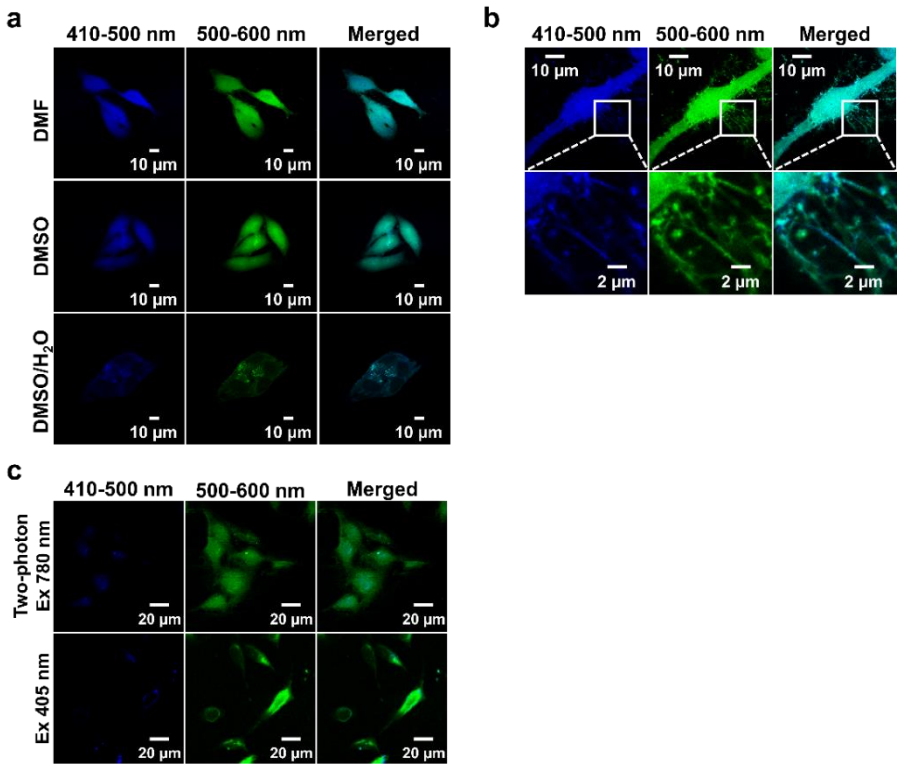


Figure 20. (a) CLSM images of *HeLa* cell samples stained by the DNFC method after 30 min in various solvents (solvents: DMF, DMSO and DMSO/H_2O). (b) CLSM images of the DNFC-stained *HeLa* cells at the focal plane near the plasma membrane (upper panel) and the magnified CLSM images of the white squares (lower panel). (c) CLSM images of the DNFC-stained *HeLa* cells using different laser sources. The upper and lower panel images were obtained by excitation with a two-photon laser (2 mW) at 780 nm and a single-photon laser (5 mW) at 405 nm, respectively. Without indication, the DNFC staining was performed in DMSO/H_2O and the CLSM images were obtained with a 405 nm-laser (5 mW).

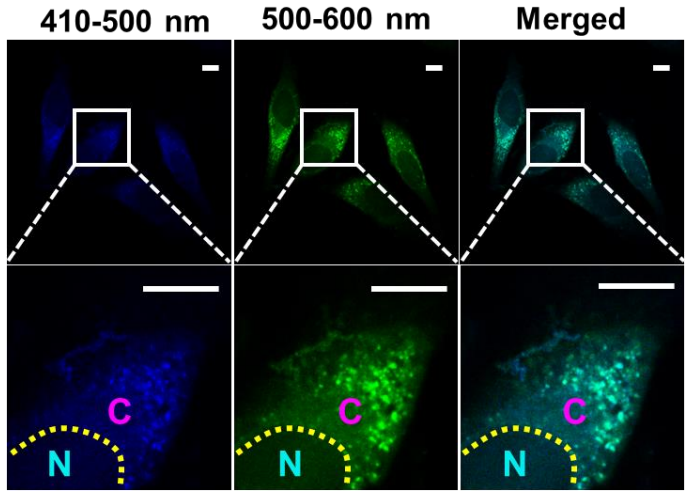


Figure 21. Magnified CLSM images of *HeLa* cells stained by the DNFC method (N: nucleus, C: cytoplasm). Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm with a laser power of 5 mW. All scale bars represent 10 μm.

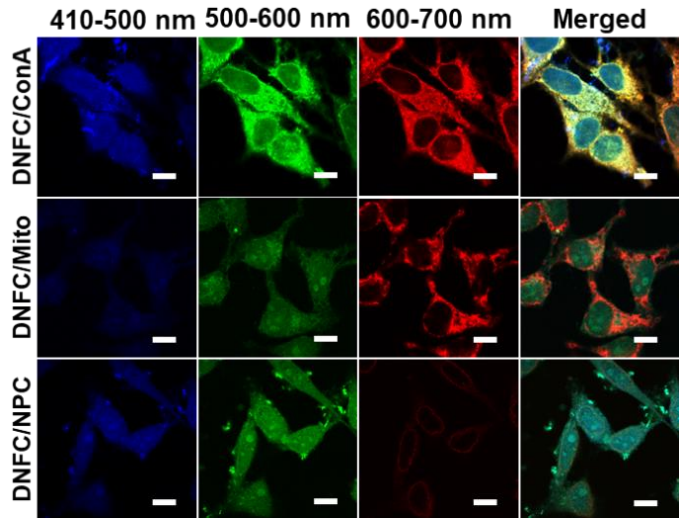


Figure 22. CLSM images of *HeLa* cells stained by the DNFC method and by each organelle-staining reagent. ER, mitochondria (Mito), and nuclear pore complex (NPC) were stained with concanavalin A (ConA), anti-TOMM20 antibody and anti-NPC antibody, which are labeled with Alexa Fluor 594 (ConA and anti-NPC) or Alexa Fluor 647 (anti-TOMM20). Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm with a laser power of 5 mW. Red channel (600-700 nm) images of ConA, TOMM20, and NPC were obtained under excitation at 543 nm with a laser power of 10 mW, 633 nm with a laser power of 5 mW and 543 nm with a laser power of 10 mW, respectively. All scale bars represent 10 μm.

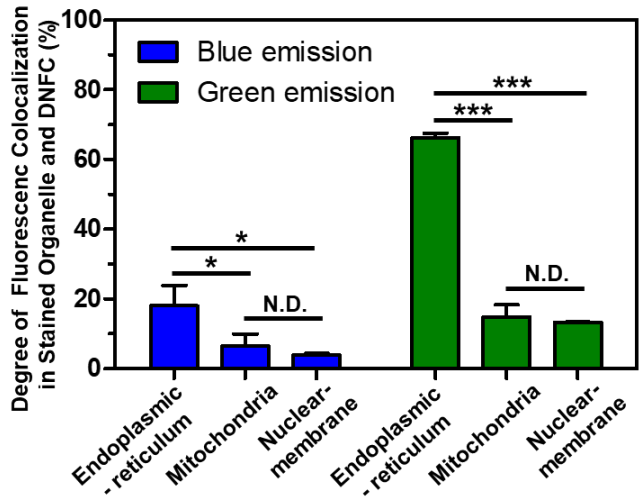


Figure 23. Colocalization ratio of blue/green fluorescence with red fluorescence. The colocalization degree was calculated from over 500 points in a CLSM image by the Imaris microscopy analysis software. Each bar represents average ± S.D. of four images (n = 4). The * and *** mean $p < 0.05$ and $p < 0.0001$, respectively (Student's t-test).

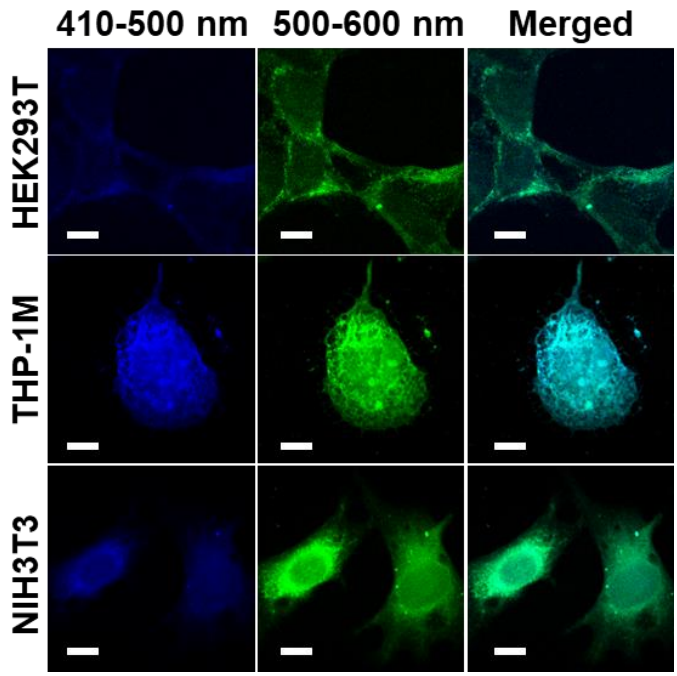


Figure 24. CLSM images of *HEK293T*, *THP-1M* and *NIH3T3* cells stained by the DNFC method. Without indication, the DNFC staining was performed for 30 min in DMSO/H_2O (70/30). Blue (410-500 nm) and green (500-600 nm) channel images were obtained under excitation at 405 nm with a laser power of 5 mW. All scale bars represent 10 μm.

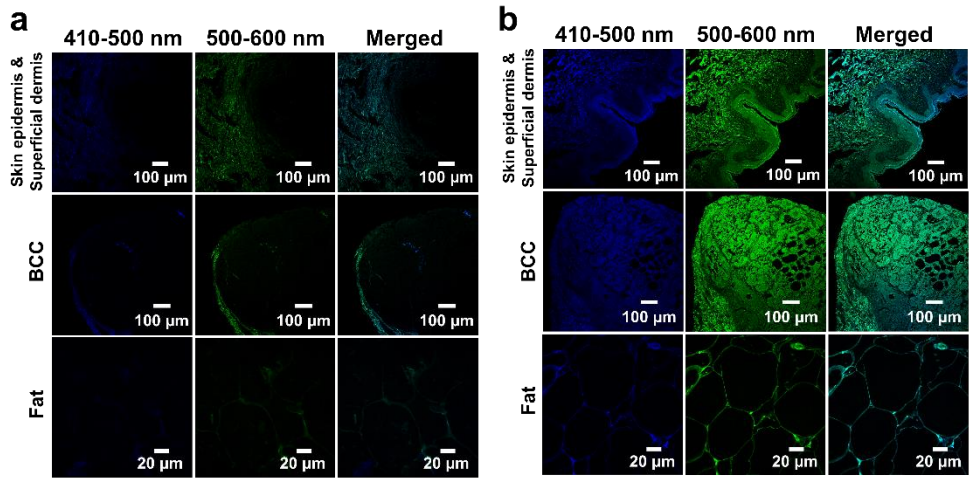


Figure 25. (a) CLSM images of a human skin/connective tissue and basal cell carcinoma (BCC) without the DNFC staining. The weak autofluorescence was observed in the image. (b) CLSM images of the DNFC-stained human skin tissue/connective tissue and BCC. All CLSM images were obtained under excitation at 405 nm. Blue (410-500 nm) and green (500-600 nm) channel images were obtained with laser power of 5 mW.

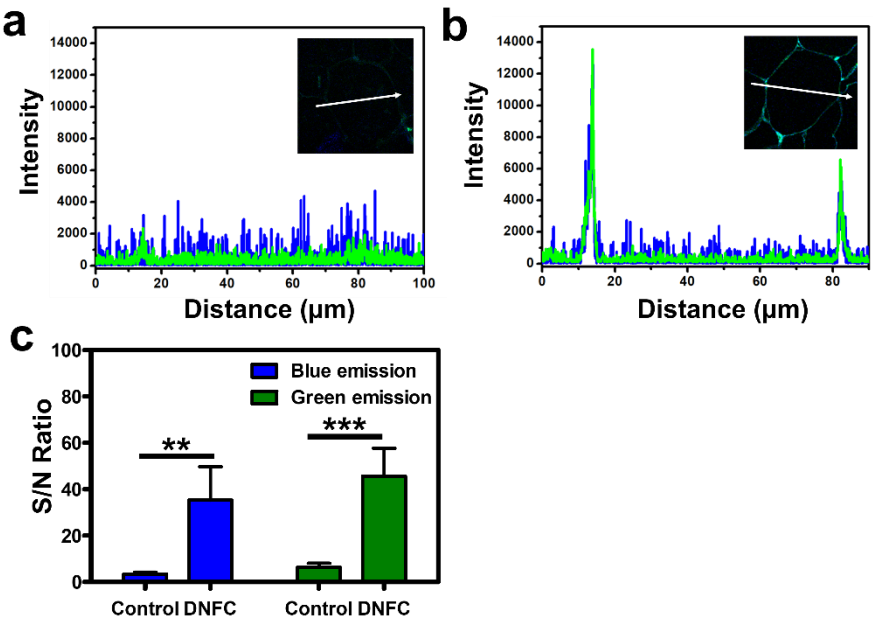


Figure 26. Line profiles indicating the distribution of the blue (blue line) and green (green line) fluorescence in the fat tissue (a) without DNFC staining and (b) with DNFC staining through the white arrow line. (c) Comparison of the signal to noise (*S/N*) ratios of the fat tissues without DNFC staining (control; autofluorescence) and with DNFC staining. The fluorescence intensities of signals and noises were analyzed by using ImageJ software. Four different tissue samples ($n = 4$) were used for the comparison. Total 100 points were compared for blue fluorescence emission and green fluorescence emission, respectively. Each bar represents the average ± S.D. The ** and *** marks mean $p < 0.01$ and $p < 0.001$, and respectively (Student's *t*-test). All CLSM images were obtained under excitation at 405 nm. Blue (410-500 nm) and green (500-600 nm) channel images were obtained with laser power of 5 mW.

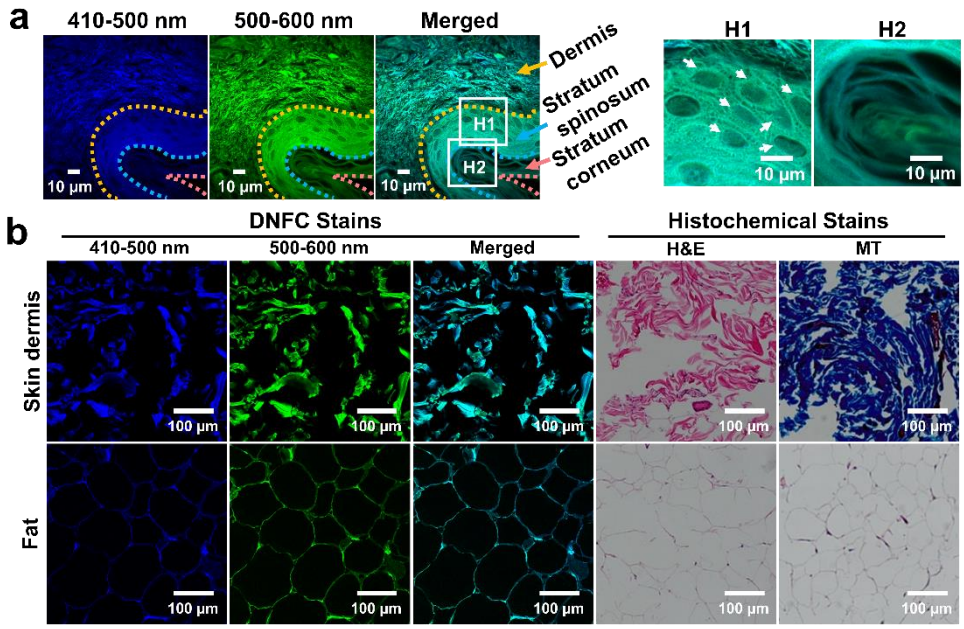


Figure 27. (a) CLSM images of a human skin/subcutaneous tissue sample strained with the DNFC method at a high magnification (H1: stratum spinosum compartment, arrows indicate polyhedral keratinocytes, H2: multiple layers of flake-like keratin in stratum corneum). (b) CLSM and optical microscopic (OM) images of skin/connective tissues stained by the DNFC and conventional histochemical methods (H&E and MT staining) at a low magnification. All DNFC staining was performed for 30 min in DMSO. The DNFC-, H&E-, and MT-stained tissue slices were obtained by serial microsectioning of tissue-paraffin blocks.

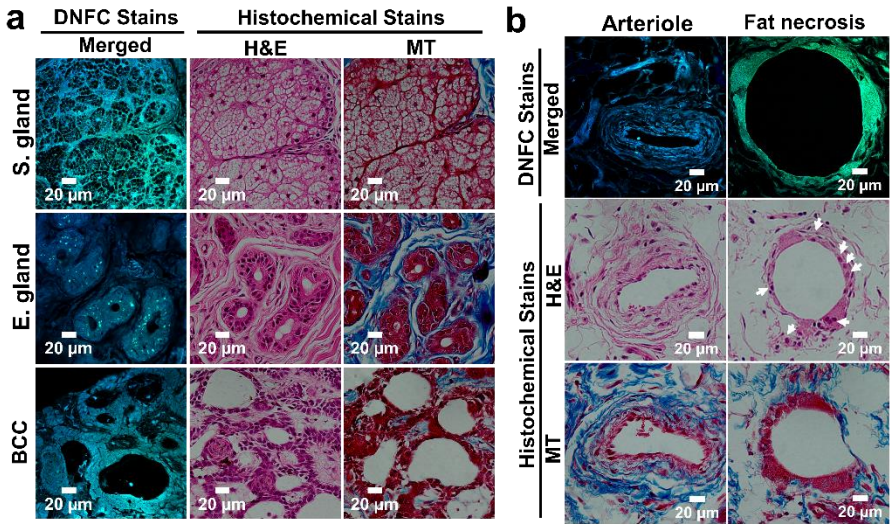


Figure 28. (a) CLSM and OM images of substructures of skin tissue samples, sebaceous glands (s. gland) and eccrine glands (e. gland), and basal cell carcinoma (BCC). (b) CLSM and OM images of substructures of subcutaneous tissues; arteriole and fat (arrows indicate histiocytes surrounding an adipocyte in fat necrosis). All DNFC staining was performed for 30 min in DMSO. The DNFC-, H&E-, and MT-stained tissue slices were obtained by serial microsectioning of tissue-paraffin blocks.

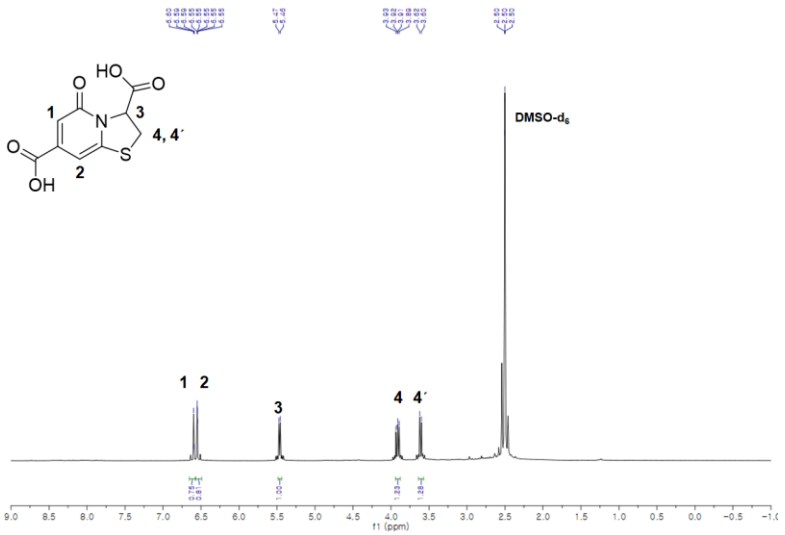


Figure 29. [1]H NMR spectrum of TPA in DMSO-d_6.

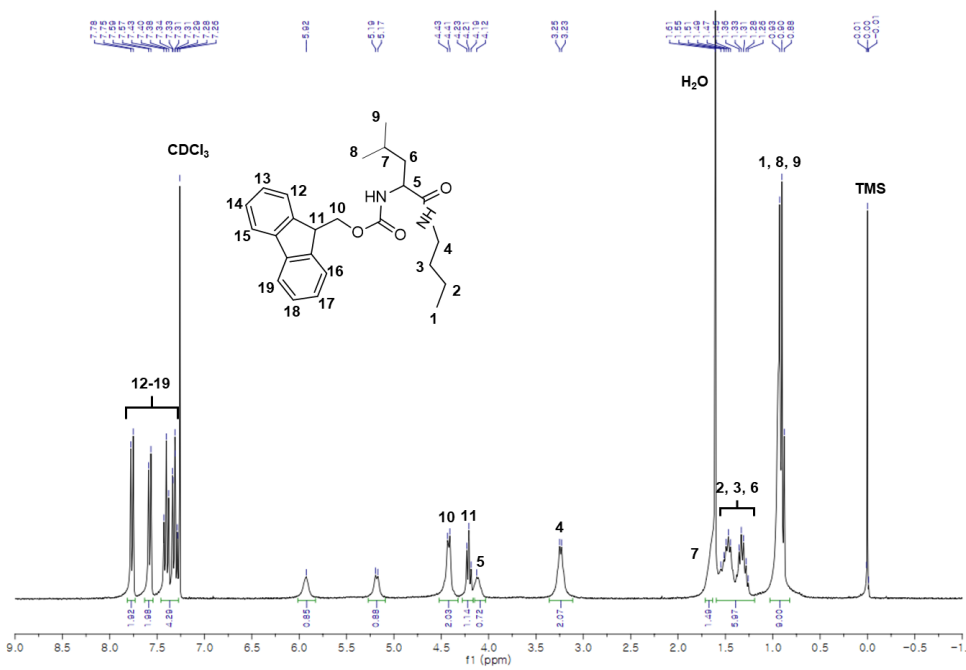


Figure 30. [1]H NMR spectrum of compound **2** in $CDCl_3$.

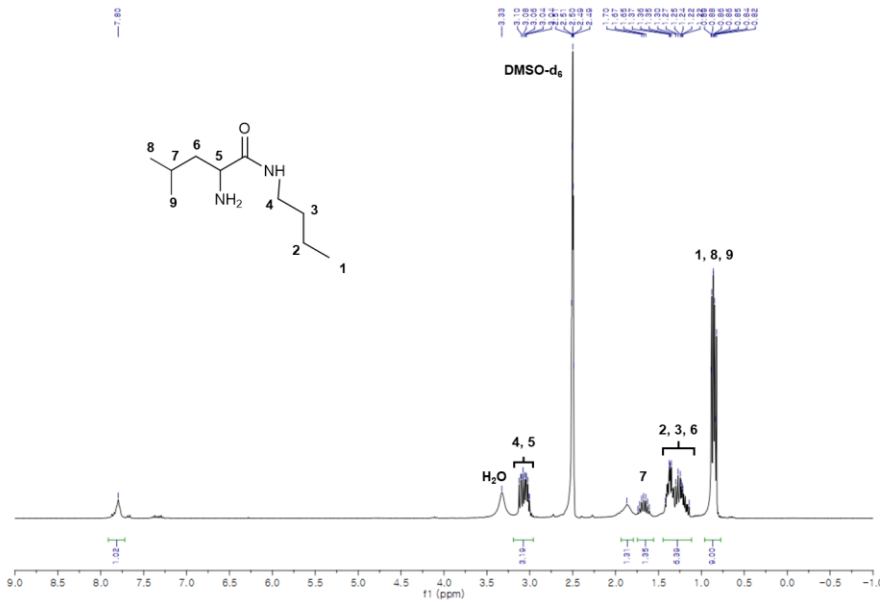


Figure 31. ^{1}H NMR spectrum of compound **3** in $CDCl_3$.

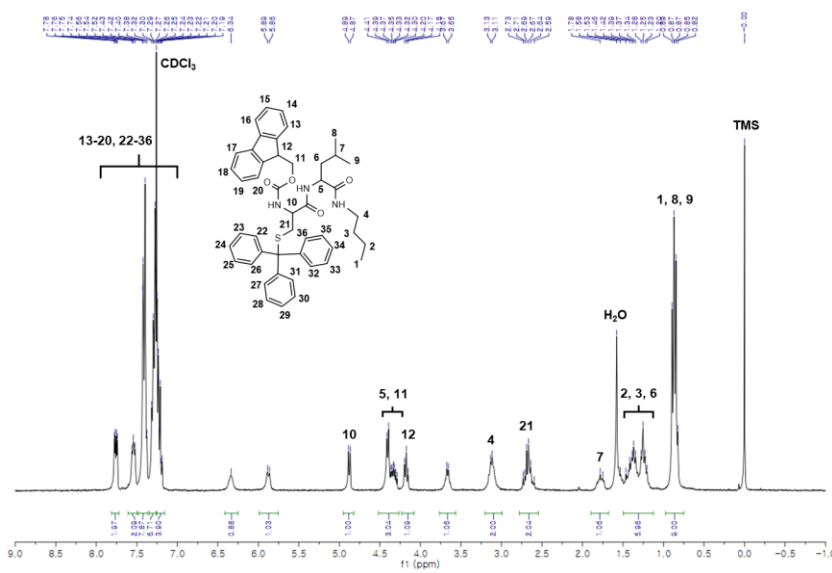


Figure 32. ^{1}H NMR spectrum of compound **4** in $CDCl_3$.

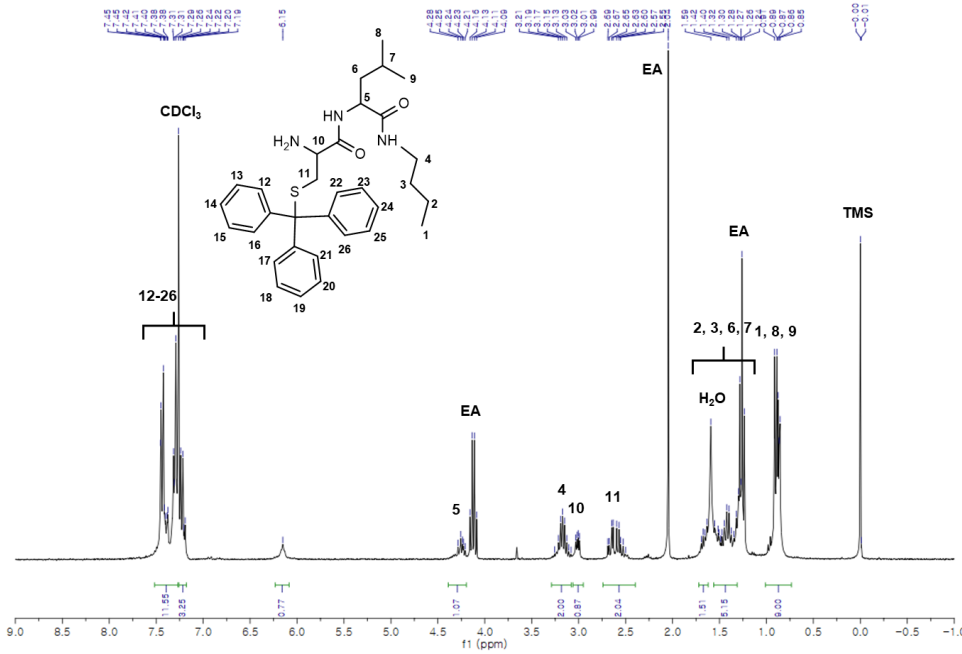


Figure 33. ^{1}H NMR spectrum of compound **5** in $CDCl_3$.

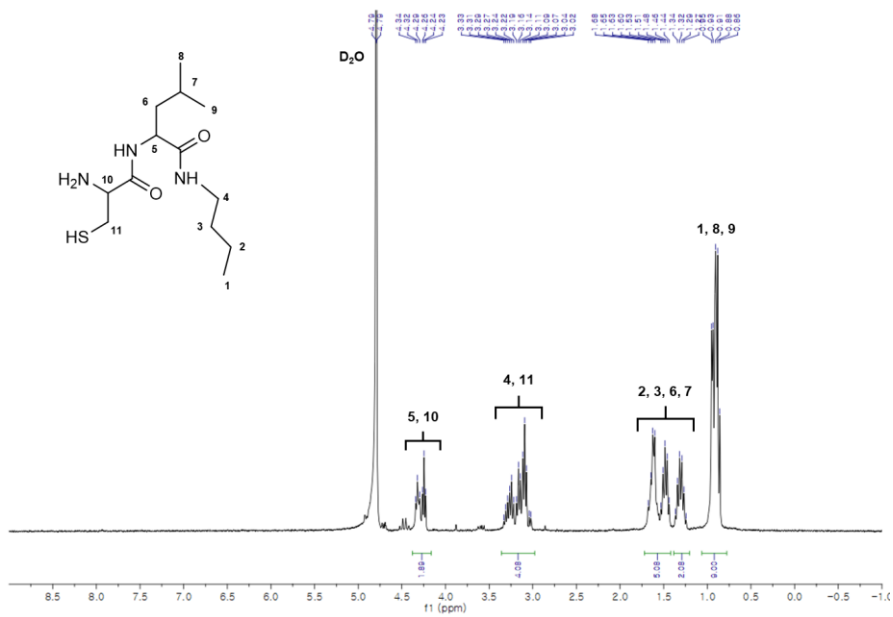


Figure 34. ^{1}H NMR spectrum of compound **6** in D_2O.

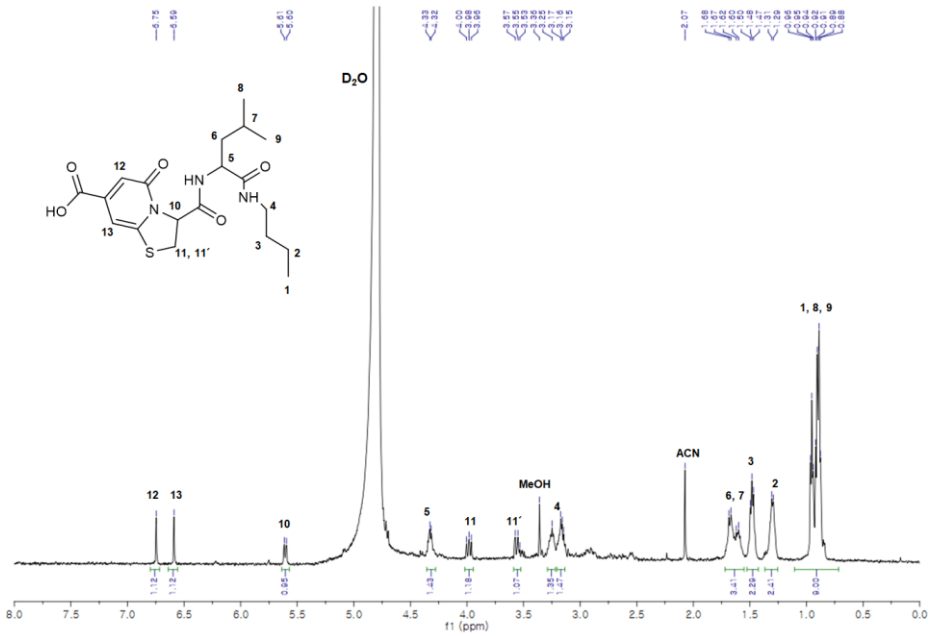


Figure 35. ^{1}H NMR spectrum of compound **7** in D_2O.

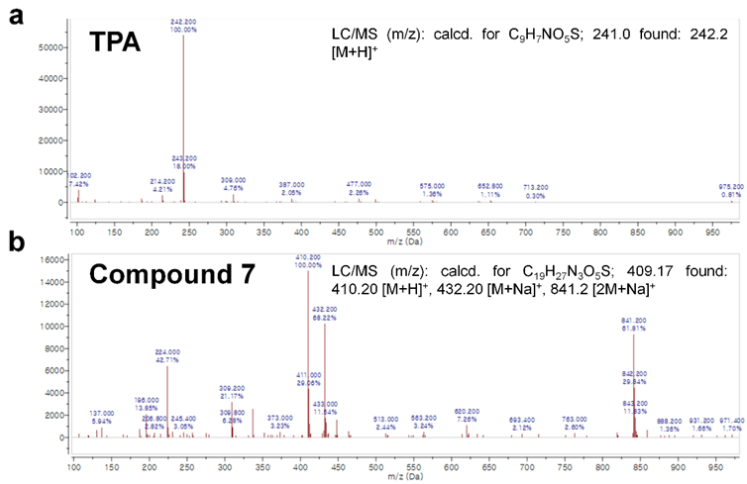


Figure 36. LC/MS spectra of TPA (a) and compound **7** (b).

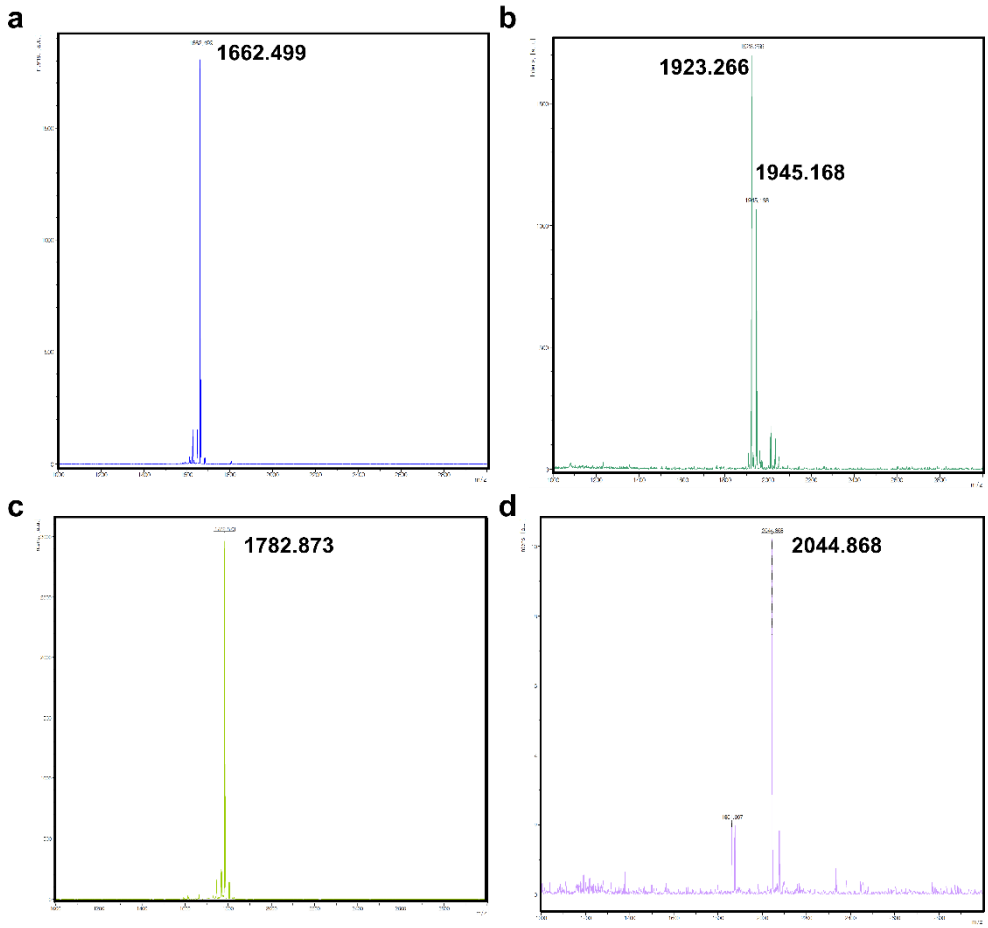


Figure 37. MALDI-TOF spectra of Cys-TAT peptide (a), Cys-LK peptide (b), TPA-TAT peptide (c) and TPA-LK peptide (d).

Table 1. Summary of DNFC-staining characteristics of cells in this research.

	Name & Origin	Appearance	Staining characteristics
Cells	*HeLa* (human cervical cancer)	Polar	Fluorescence preferentially developed in the cytoplasm (ER) rather than in the nuclei green intensity > blue intensity
	HEK293T (human embryonic kidney)	Multipolar	Fluorescence preferentially developed in the cytoplasm rather than in the nuclei Lower fluorescence intensity than other cells green intensity > blue intensity
	THP-1M (human monocyte)	Irregular spherical	Much stronger blue fluorescence than other cells
	NIH3T3 (mouse fibroblast)	Multipolar	Fluorescence preferentially developed in the cytoplasm rather than in the nuclei green intensity > blue intensity

Table 2. Summary of DNFC-staining characteristics of tissues in this research.

	Name & Origin	Appearance	Staining characteristics
Human tissues	Skin & subcutaneous fat (keloid of earlobe)	Epidermis (stratum corneum and stratum spinosum layers) Dermis (collagen and elastic fibre bundle) Subcutaneous fat (adipocyte)	Keratin layers (stratum corneum), extracellular matrix of keratinocytes (stratum spinosum), fibre bundles (dermis) and extracellular matrix/cytoplasm of adipocytes (subcutaneous fat) strongly stained No staining in the lipid droplets
	Basal cell carcinoma (facial skin)	Subcutaneous substructures (sebaceous and eccrine glands) Tumor cells and stroma	Mesh-like structure (sebaceous gland), bright dot-like structures (eccrine gland) and tumor cells strongly stained No staining in the stroma
	Dermato-fibrosarcoma protuberance	Arteriole (thick wall and lumen) Fat necrosis (histiocytes and multinucleated giant cells)	Thick arteriole wall, histiocytes and giant cells strongly stained No staining in fat

PART II. Chemoselective Tyrosine Bioconjugation Through Sulfate Click Reaction

1. Abstract

A novel and selective tyrosine functionalization strategy through SuFEx (sulfur fluoride exchange) chemistry is presented. In this approach, free tyrosine (Tyr) reacts selectively with aryl fluorosulfate in the presence of various nucleophilic amino acid residues in bio-tolerable conditions. Chemoselectivity of this unique SuFEx reaction was confirmed in amino acid, peptide, and protein conjugations. The functions of peptides and proteins were well preserved as demonstrated from the Tyr-specific modification of cell-penetrating peptide and erythropoietin. This method is well-suited for residue-specific modification of native proteins, and thus would expand the versatility of bio-conjugation in protein chemistry.

2. Introductions

Site-specific functionalization of biomolecules is of surmounting importance, since non-selective functionalization generates heterogeneous mixtures with often unpredictable activities.[1] Apart from purely biological methods such as enzymatic reactions,[2] most biochemical functionalization methods are categorized into two types; 1) bioorthogonal functionalization of a genetically modified site (an unnatural amino acid (UAA) residue)[3] and 2) selective chemical reaction of a natural amino acid residue (Figure 1).[4] As an example of the former, azidophenylalanine (AzF) has been inserted for the functionalization handle through the use of expanded genetic codes.[3] AzF can react specifically with an alkyne-containing molecule via cycloaddition reaction. However, the incorporation of UAAs is operationally complex with often moderate expression yield.[5]

There are ample examples of readily accessible natural protein functionalization methods in the second category. Lysine (Lys)-succinimide and cysteine (Cys)-maleimide coupling reactions have been widely used. Since most proteins expose several Lys residues on the surface, the Lys-succinimide coupling reactions produce a mixture of products.[6] Cys residues are scarce on the surface of most proteins and readily oxidized to disulfides.[7] In addition, Lys can also react with a maleimide group in basic conditions.[8] More recently, alternative strategies for selective amino acid tagging have been developed. Francis' group reported an N-terminus modification strategy through pyridoxal phosphate (PLP) mediated transamination and subsequent oxime formation.[9, 10] Toste and Chang's group discovered a selective methionine (Met) bioconjugation using oxaziridine reagents.[11] Functionalization on only one *N*-terminus or a rare amino acid such as Met appears to be attractive for the maximization of selectivity;[12] however, the rarity of the functional handle may be a drawback for efficient bioconjugation. Therefore, development of new strategies for selective functionalization on specific

amino acid residues is still in need.

Recently, Sharpless developed a sulfate click reaction between an aryl silyl ether and an aryl fluorosulfate (fosylate) using a catalytic amount of a strong base such as 1,8-diazabicyclo[5.4.0]undec-7ene (DBU), named SuFEx reaction (Figure 2A).[13] Though extremely useful, its use in direct tyrosine (Tyr) bioconjugation has been scarce since silylated Tyr residue was presumed to be required. Exceptionally, a report from Kelly, Sharpless and co-workers described the sulfate click reaction of a free Tyr embedded within an active site of LBP (lipid binding protein).[14] In this case a nearby arginine (Arg) was presumed to assist the reaction within the active site. Previously, we had observed that DBU can promote the desilylation of various aryl silyl ethers in polar environment (Figure 2B).[15] Therefore, in the SuFEx reaction the initial deprotection of an aryl silyl ether mediated by DBU would generate an aryloxide anion, which in turn reacts with the fosylate producing a fluoride anion, hence the completion of the catalytic cycle (Figure 3). This obviates the need of an aryl silyl ether in the modified SuFEx reaction, where a stoichiometric amount of base is used. To confirm this hypothesis, we checked the reactivity of *p*-cresol with phenyl fosylate in presence of various bases (Figure 4). With tetramethylguanidine (TMG) and DBU, the desired biaryl sulfate was synthesized quantitatively without the need of pre-silylation of *p*-cresol. Based on these results, we propose a novel chemoselective conjugation strategy that utilizes the coupling reaction of a Tyr residue in proteins through the modified SuFEx reaction utilizing free Tyr.

There have been a few reports on Tyr-selective bioconjugation using Mannich or ene-type reactions.[16, 17] However, they require excess reagents for full conversion. Our strategy was also supported by a previous report that a fosylate-based bioprobe could form a covalent bond on a Tyr residue at a binding site of LBP,[14] where a nearby Arg residue appeared to assist the reaction between the bioprobe and Tyr.

While the previous study reported the reaction between Tyr only at the binding site of the protein and fosylate probe,[14, 18] we envisioned that the application of SuFEx reaction could be expanded to general Tyr functionalization on protein surface.

3. Experimental Section

3.1. Materials

N-α-Fmoc protected L-amino acids, Rink Amide MBHA resin (0.45 mmol/g loading), and (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP) were purchased from BeadTech. *N*,*N*-Dimethylformamide (DMF), dimethyl sulfoxide (DMSO), 1,2-dichloromethane (DCM), acetonitrile (ACN), n-hexane and diethyl ether were purchased from Samchun. Chemicals including Rhodamine, 4-hydroxybenzoic acid, *p*-cresol, DIPEA, triisopropylsilane (TIS), trifluoroacetic acid, and piperidine were purchased from TCI, Alfa Aesar and Sigma-Aldrich. Hoechst 33342 solution and cell counting kit-8 (CCK-8) were purchased from Dojindo. Dulbecco's modified eagle's medium (DMEM), Dulbecco's phosphate buffered saline (DPBS) and fetal bovine serum (FBS) were purchased from WELGENE. 0.05% Trypsin-EDTA (1×) was purchased from Gibco. Erythropoietin (EPO; human) was purchased from GenScript. Centrifugal Filter Units were purchased from Milipore.

3.2. Instruments

The ^{1}H and ^{13}C NMR-spectra were measured with a Varian/Oxford As-500 (500 MHz), an Agilent 400-MR DD2 Magnetic Resonance System (400 MHz) and a Bruker DPX-300 (300 MHz) spectrophotometers. Chemical shifts were measured as part per million (δ values) from tetramethylsilane as an internal standard at probe temperature in CD_3OD or $CDCl_3$ or DMSO-D_6 for neutral compounds. Reactions that needed anhydrous conditions were carried out in flame-dried glassware under positive pressure of dry Ar using standard Schlenk line techniques. Evaporation of solvents was performed at reduced pressure using a rotary evaporator. Reactions of small molecules were monitored with thin layer chromatography (TLC), which was

performed using silica gel 60F254 coated on aluminum sheet (E. Merck, Art.5554). Chromatogram was visualized by UV-lamp (Vilber Lournat, VL-4LC) and/or colorized with following solutions: (a) 20% ethanolic phosphomolybdic acid (PMA), (b) potassium permanganate solution, and (c) 2% ninhydrin ethanolic solution. Reactions of polymers, peptides, and proteins were monitored with MALDI/TOF mass analyzer (Bruker, DE/microflex LT and ultraflextreme). Column chromatography was performed on silica gel (Merck. 7734 or 9385 Kiesel gel 60), and eluent was mentioned in each procedure. HPLC analyses were carried out on an HP1100 system Agilent (Santa Clara, CA, USA) or LC-20 series SHIMADZU, composed of auto sampler, quaternary pump, photodiode array detector (DAD) and HP Chemstation software. The separation was carried out on a C18 Vydac 218TP54 column 250 x 4.6 mm i.d. (5 μm particle size), or C18 Zorbax (5 μm, 9.4x250 mm) with 0.1% TFA in water (A), acetonitrile (B), as a mobile phase at a flow rate of 1 mL/min at 20°C. 10 K Amicon® Ultra-0.5 centrifugal filter units (Millipore) were used for ultracentrification, and the protein concentrations were measured by NanoDropTM2000 spectrophotometer.

3.3. Synthesis of phenyl *p*-tolyl sulfate (1)

A mixture of *p*-cresol (21.6 mg, 0.200 mmol) and phenol fosylate (35.2 mg, 0.200 mmol) was stirred in DCM (2.00 mL) with TMG (25.2 μL, 0.200 mmol) at room temperature for 1 h. The reaction mixture was dried, concentrated under reduced pressure and then purified with column chromatography (EA/HEX). The residue was obtained as a white solid (51.8 mg, 98.0%). ^{1}H NMR ($CDCl_3$, 400 MHz): 7.42 (m, 2H), 7.33 (m, 3H), 7.19 (s, 4H), 2.36 (s, 3H). ^{13}C NMR ($CDCl_3$, 100 MHz): 150.5, 148.3, 137.6, 130.5, 130.0, 127.5, 121.0, 120.8, 20.9. HRMS (ESI) *m/z*: Anal. calcd. For $[M+Na]^+$ $C_{13}H_{12}O_4SNa$: 287.03; found: 287.0352.

3.4. Synthesis of *N*-Boc tyramine (2)

N-Boc tyramine was synthesized through a procedure described in the literature.[19]

3.5. Synthesis of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (3)

A mixture of compound **2** (0.130 g, 0.550 mmol) and triethylamine (0.240 mL, 1.69 mmol) was stirred in DCM (2.75 mL) under SO_2F_2 atmosphere at room temperature. After 5 h, the mixture was poured into distilled water (20.0 mL) and extracted with DCM (20.0 mL) 3 times. The organic layer was dried over $MgSO_4$, and then was purified with column chromatography (EA/Hex). A pale pink solid (145 mg, 82.9%) was obtained. ^{1}H NMR ($CDCl_3$, 400 MHz): 7.26 (s, 4H), 4.53 (br s, 1H), 3.36 (d, J = 8.0 Hz, 2H), 2.82 (t, J = 8.0 Hz, 2H), 1.41 (s, 9H). ^{13}C NMR ($CDCl_3$, 100 MHz): 155.8, 148.6, 140.0, 130.7, 120.8, 79.4, 41.5, 35.7, 28.3. ^{19}F NMR ($CDCl_3$, 400 MHz): 37.26. LRMS (ESI) *m/z*: Anal. calcd. For $[M+Na]^+$ $C_{13}H_{18}FNNaO_5S$: 342.08; found: 342.2.

3.6. Synthesis of tetraethylrhodamine succinimidyl ester (4)

N-Boc tyramine was synthesized through a procedure described in the literature.[20]

3.7. Synthesis of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl) carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (5)

Compound **3** (46.4 mg, 0.150 mmol) was stirred in 10% TFA in DCM (1.00 mL) for 1 h. The mixture was poured into 1 N NaOH solution (50.0 mL) in an ice bath and extracted with DCM (30.0 mL) 3 times. The organic layer was dried over $MgSO_4$, filtered and concentrated under reduced pressure. To the residue in anhydrous DCM (2.00 mL) were added compound **4** (40.0 mg, 0.070 mmol) and TEA (20.0 μL, 0.140 mmol) and the mixture was stirred in for 2 h. The reaction mixture was then poured into brine (10.0 mL) and extracted with DCM (10.0 mL) 3 times. The organic layer was dried over $MgSO_4$, and then was purified with column chromatography (EA/Hex). A yellow oil (11 mg, 24.3%) was obtained. ^{1}H NMR ($CDCl_3$, 500 MHz): 7.92 (t, *J* = 5.0 Hz, 1H), 7.46 (dd, *J* = 5.0 Hz, 2H), 7.13-7.06 (m, 5H), 6.41 (s, 3H), 6.39 (s, 1H), 6.24 (dd, *J* = 2.5, 7.5 Hz, 2H), 3.33 (q, *J* = 6.7 Hz, 8H), 3.29 (t, *J* = 7.5 Hz, 2H), 2.53 (t, *J* = 7.5 Hz, 2H), 1.17 (t, *J* = 7.5 Hz, 12H). ^{13}C NMR (CDCl3, 125 MHz): 167.8, 153.5, 153.1, 148.8, 148.3, 140.5, 132.4, 131.5, 130.6, 128.9, 128.1, 123.8, 122.7, 120.5, 108.1, 105.6, 97.6, 65.0, 44.3, 41.7, 33.9, 29.7, 12.6. ^{19}F NMR ($CDCl_3$, 300 MHz): 37.18. HRMS (ESI) *m/z*: Anal. calcd. For $[M]^+$ $C_{36}H_{39}FN_3O_5S$: 644.26; found: 644.2593.

3.8. Synthesis of 4-((Fuorosulfonyl)oxy)benzoic acid (6)

A mixture of 4-hydroxy benzoic acid (276 mg, 2.00 mmol), and TEA (1.10 mL, 8.00 mmol) was stirred under sulfuryl fluoride atmosphere in DCM (5.00 mL) for 12 h. The mixture was poured into 1 N HCl solution (10.0 mL) and extracted with DCM (10.0 mL) 3 times. The organic layer was dried over $MgSO_4$, concentrated under reduced pressure and then was purified with column chromatography

(EA/Hex). A yellow solid (114 mg, 26.0%) was obtained. ^{1}H NMR ($CDCl_3$, 400 MHz): 10.19 (br s, 1H), 8.19 (d, J = 8.0 Hz, 2H), 7.42 (d, J = 12.0 Hz, 2H). ^{13}C NMR ($CDCl_3$, 125 MHz): 169.7, 153.4, 132.7, 128.7, 121.1. ^{19}F NMR ($CDCl_3$, 400 MHz): 38.77. LRMS (ESI) *m/z*: Anal. calcd. For $[M-H]^-$ $C_7H_4FO_5S$: 218.98; found: 218.95.

3.9. Synthesis of 2,5-Dioxopyrrolidin-1-yl 4-((fluorosulfonyl)oxy)benzoate (7)

A mixture of compound **6** (113 mg, 0.510 mmol), *N*-hydroxysuccinimide (226 mg, 2.00 mmol), and *N*,*N*'-dicyclohexylcarbodiimide (117 mg, 0.570 mmol) was stirred in anhydrous THF (5.00 mL) for 12 h. The mixture was concentrated under reduced pressure and purified with column chromatography (EA/Hex). A white solid (150mg, 92.2%) was obtained. ^{1}H NMR ($CDCl_3$, 300 MHz): 8.31 (d, J = 9.0 Hz, 2H), 7.54 (d, J = 9.0 Hz, 2H), 2.94 (s, 3H). ^{13}C NMR ($CDCl_3$, 75 MHz): 169.0, 160.4, 154.0, 133.1, 128.6, 121.7, 25.7. ^{19}F NMR ($CDCl_3$, 300 MHz): 39.32. LRMS (ESI) *m/z*: Anal. calcd. For $[M+K]^+$ $C_{11}H_8FNO_7SK$: 356.0; found: 356.2.

3.10. Synthesis of methoxy poly ethylene glycol tosylate (8)

Methoxy poly ethylene glycol-2000 (MPEG-2000, 2.30 g, 1.15 mmol), TsCl (0.400 g, 2.10 mmol), and TEA (0.400 mL, 2.90 mmol) were stirred in DCM (2.00 mL) at 50°C for overnight. The mixture was poured into 1 N NaOH solution (40.0 mL) and extracted with DCM (10.0 mL) 3 times. The organic layer was washed with 1 N HCl solution (20.0 mL) 3 times, then was dried over $MgSO_4$ and evaporated. The residue was dissolved in a small amount of DCM and triturated with ethyl ether (15.00 mL). A white solid (1.56 g, 63.2%) was obtained. ^{1}H NMR ($CDCl_3$, 400 MHz): 7.79 (d, J = 8.0 Hz, 2H), 7.34 (d, J = 8.0 Hz, 2H), 4.16-3.45 (m, ~180H), 3.37 (s, 3H), 2.44 (s, 3H).

3.11. Synthesis of methoxy poly ethylene glycol azide (9)

A mixture of compound **8** (400 mg, 0.190 mmol), sodium azide (24.2 mg, 0.370 mmol), and sodium bicarbonate (23.5 mg, 0.280 mmol) was stirred in DMF (1.0 mL) at 120°C for 12 h. The solvent was completely evaporated, and the residue dissolved in a small amount of DCM and triturated with ethyl ether (5.0 mL). A yellow solid (363 mg, 96.4%) was obtained.

3.12. Synthesis of methoxy poly ethylene glycol amine (10)

A mixture of compound **9** (363 mg, 0.180 mmol) and triphenylphosphine (70.5 mg, 0.270 mmol) was stirred in anhydrous THF (1.00 mL) for 12 h. Water (100 μL) was added to the mixture and it was stirred for 12 h. The solvent was evaporated, and the residue was solubilized in 0.5 N NaOH solution (10.0 mL) and extracted with DCM (10.0 mL) 3 times. The organic layer was dried over $MgSO_4$ and evaporated. The residue dissolved in a small amount of DCM and triturated with ethyl ether (5.00 mL). The precipitate was washed with hexane to yield a white solid (278 mg, 77.3%). ^{1}H NMR ($CDCl_3$, 400 MHz): 3.83-3.49 (m, ~180H), 3.38 (s, 3H), 2.86 (t, J = 4.0 Hz, 2H).

3.13. Synthesis of methoxy poly ethylene glycol amine (11)

A mixture of compound **7** (17.5 mg, 0.055 mmol), **10** (100 mg, 0.05 mmol), and TEA (14 μL, 0.1 mmol) was stirred in anhydrous THF. After 12 h, the mixture was poured into 1 N HCl solution and extracted with DCM. The organic layer was dried over $MgSO_4$ and evaporated. The residue was solubilized with small amount of DCM and precipitated with ethyl ether. Yellow solid (102 mg, 92.7%) was obtained.

^{1}H NMR ($CDCl_3$, 400 MHz): 7.99 (d, J = 8.0 Hz, 2H), 7.41 (d, J = 8.0 Hz, 2H), 3.83-3.45 (m, ~180H), 3.38 (s, 3H). ^{19}F NMR ($CDCl_3$, 300 MHz): 38.36.

3.14. Synthesis of (2S,2'S)-3,3'-((Sulfonylbis(oxy))bis(4,1-phenylene))bis(2-aminopropanoic acid) (12)

N-Boc-L-tyrosine (118 mg, 0.420 mmol), sulfonyl diimidazole (40.0 mg, 0.200 mmol), and cesium carbonate (391 mg, 1.20 mmol) were stirred in DMF (2.00 mL) at 60°C. After 12 h, the mixture was poured into 1 N HCl (20.0 mL) and extracted with DCM (10.0 mL) 3 times. The organic layer was dried over $MgSO_4$ and evaporated. The residue was solubilized with 30% TFA solution in DCM (2.00 mL) and stirred for 2 h. TFA and solvent were evaporated with rotary, and the residue was purified with reverse phase HPLC (water/acetonitrile). White solid (29.5 mg, 34.8% for two steps) was obtained. ^{1}H NMR (CD_3OD, 400 MHz): 7.42 (d, J = 8.0 Hz, 4H), 7.32 (d, J = 8.0 Hz, 4H), 4.13 (t, J = 6.0 Hz, 2H), 3.33-3.30 (m, 2H), 3.18 (dd, J = 6.0, 12.0 Hz, 2H), 2.13 (s, 2H). HRMS (ESI) *m/z*: Anal. calcd. For $[M+H]^+$ $C_{18}H_{21}N_2O_8S$: 425.10; found: 425.1016.

3.15. Synthesis of TAT 47-57 peptide

Synthesis of TAT peptide (sequence: YGRKKRRQRRR) was performed by using microwave heating in Fmoc-based solid-phase peptide synthesis with Rink Amide MBHA resin. Peptides were synthesized in 90 μmol scale. The typical conditions[21] for microwave-assisted peptide synthesis were used: Fmoc-deprotection with 20% piperidine at 75°C (50 W) for 20 min and peptide coupling with PyBOP and DIPEA at 75°C (25 W) for 20 min.

First of all, Rink Amide resins (200 mg, 0.090 mmol, 0.45 mmol/g loading)

were deprotected with 20% piperidine in DMF (2 mL) under N_2 bubbling, and then a mixture of the resin, Fmoc-Arg (Pbf)-OH (350.3 mg, 0.540 mmol) were stirred with PyBOP (281 mg, 0.540 mmol), and DIPEA (200 μL, 1.18 mmol) under N_2 atmosphere. Reaction byproducts were removed with DMF and DCM several times. After the coupling step, Fmoc-protected amino acid was deprotected with 20% piperidine in DMF, and the reaction byproducts were removed with DMF and DCM several times. The coupling and deprotection steps were repeated with different Fmoc-protected amino acids sequentially until the deprotection of the last Fmoc-protected amino acids was complete.

Cleavage of the peptide from resins was carried out by treatment with TFA/TIS/water (total volume 2.00 mL, v/v=95:2.5:2.5) for 2 h at room temperature. The cleaved resin was then separated by filtration and further washed with TFA. The separated peptide solution was concentrated by a nitrogen gas. The synthesized peptide was precipitated in a mixture of n-hexane and diethyl ether (v/v=50:50). The resulting suspension was centrifuged at 4,000 rpm for 15 min. After the supernatant was decanted, the pellet was dissolved in DMF and was purified with HPLC. The peptide was lyophilized affording a white powder. LRMS (MALDI) *m/z*: Anal. calcd. for $[M+H]^+$ $C_{64}H_{120}N_{33}O_{13}$: 1559.0; found: 1560.0. The HPLC chromatogram of TAT peptide indicates 93% purity (absorbance was measured at 220 nm) (Figure 33A).

3.16. Synthesis of Rho-TAT peptide

The purified TAT peptide (2.00 mg, 0.810 μmol) and TMG (2.00 μL, 16.2 μmol) were dissolved in DMSO (0.8 mL), and then compound **5** (1.00 mg, 1.62 μmol) was added. The mixture was stirred at room temperature for 3 h. The reaction mixture was purified with HPLC at a flow rate of 3 mL/min. The peptide was lyophilized affording a purple colored powder. LRMS (MALDI) *m/z*: Anal. calcd. for [M]+

$C_{100}H_{157}N_{36}O_{18}S$: 2182.2; found: 2183.3. The HPLC chromatogram of Rho-TAT peptide indicates 99% purity (absorbance was measured at 220 nm) (Figure 33B).

3.17. Synthesis of PEG-rhEPO

50.0 μL ($4.63x10^{-4}$ μmol) of rhEPO stock solution (25.0 μg/100 μL) was dissolved in a Tris buffer (50 mM, pH 8.0) (450 μL) along with 10.0 μL ($4.63x10^{-2}$ μmol) of TMG stock solution (0.530 mg/mL) and 10.0 μL ($2.32x10^{-3}$ μmol) of fosylate PEG (0.510 mg/mL). After 3 h of reaction, the crude mixture was purified by ultracentrification at 14,000 rpm for 15 min (×5). Then, the buffer was exchanged into DPBS by using Amicon Ultra-0.5 devices (×3). LRMS (MALDI) *m/z*: Anal. calcd. for $[M+H]^+$: 28914.9; found: 28260.2.

3.18. Tryptic digestion of proteins

rhEPO and PEG-rhEPO were subjected to trypsin digestion through a modified method.[22] Briefly, 12.5 μg of rhEPO and PEG-rhEPO were dissolved in 30.0 μL of DPBS. Trypsin protease at an enzyme to substrate ratio of 1:20 (w/w) was added and the mixture was incubated at 37°C for 3 h.

3.19. Confocal laser scanning microscopy (CLSM) observation of Rho-TAT peptide

HeLa cells (human cervical cancer cells) (10,000 cells/well) were cultured in 180 μL of DMEM supplemented with 10% FBS on confocal dishes (SPL Lifesciences, South Korea) at 37°C in a 5% CO_2 incubator. After 24 h, the medium was exchanged with fresh DMEM supplemented with 2% FBS. Then, Rho-TAT

peptide in deionized water (final concentration: 10 μM) were added to the dish and the cells were further incubated for 12 h. Cells were washed 3 times with fresh medium for the removal of peptides outside the cells, and fresh medium was added. Nucleus was stained with a Hoechst 33342 solution (final concentration: 4 μM) for 10 min prior to the CLSM imaging. After further washing with fresh medium (×3), CLSM images were acquired by using LSM 700 (Carl Zeiss). A band pass channel (420-480 nm, blue signal) and a long-wavelength pass channel (560 nm, red signal) for Hoechst 33342 and rhodamine were used for the imaging, respectively.

3.20. *In vitro* cytotoxicity of diaryl sulfate compound

The *in vitro* cytotoxicity of diaryl sulfate bond was evaluated on the *HeLa* cells by the cell counting kit-8 (CCK-8) viability tests. Briefly, *HeLa* cells were seeded in 96-well tissue culture dishes at 5,000 cells/well and incubated in 100 μL of DMEM containing 10% FBS for 24 h. After the medium was replaced with 90 μL of fresh medium, 10 μL of various concentrations of **12** (final 0.5% DMSO) were added to the wells and the cells were further incubated for 24 h and 48 h. Cells were washed 3 times with fresh medium for the removal of the sample outside cells, and 100 μL of fresh medium containing 10% CCK-8 was added to each well. After incubation for 2 h at 37°C, absorbance at 450 nm was measured by a microplate reader (Molecular Devices Co., Menlo Park, CA, USA). The background signal of CCK-8 alone was subtracted from all samples. Cell viability was evaluated as a percentage of absorbance relative to the control of untreated cells.

$$Cell\ \ Viability\ \ (\%) = \frac{Absorbance_{treated} - Absorbance_{background}}{Absorbance_{untreated} - Absorbance_{background}} \times 100$$

3.21. *In vivo* activity of rhEPO and PEG-rhEPO

For effects of hematopoietic activity of rhEPO and PEG-rhEPO, normal male Balb/c mice were purchased from Orient bio inc. (South Korea). The mice weighed 23-25 g and were group-housed (3-4 per cage). Food and water were available *ad libitum* and the room had a 12 h light/dark cycle under a conventional animal experiment system at KPC, South Korea. The Institutional Animal Care and Use Committee at KPC approved all procedures.

Briefly, normal male Balb/c (7/group) mice were intravenously administrated with 50 μL of each sample (0.16 μM in DPBS) every third day during 2 weeks. The control group was administrated with an equivalent amount of DPBS. A total of 21 mice were used in this experiment. Blood samples were collected in every 3 days for evaluating the effect on hematogeniety. Hematocrit was evaluated by measurement of packed cell volume immediately after blood collection. Delta hematocrit (ΔHematocrit) was determined by the difference between the initial hematocrit (day 0) and the hematocrit at each time point for each mouse.

$$\Delta\text{Hematocrit } (\%) = Hematocrit\ (\%)_{day} - Hematocrit\ (\%)_{day\ 0}$$

4. Results and Discussion

4.1. Comparison of SuFEx reactivity of various model nucleophiles representing amino acids

As the first step of developing the protein conjugation strategy, we studied the chemoselectivity of our modified SuFEx on a simple conjugating molecule, phenyl fosylate, towards various model counterpart molecules corresponding to representative nucleophilic amino acid residues (Table 1). Propanethiol was selected as a model nucleophile of Cys instead of gaseous methanethiol. TMG was adopted as a base for the reaction, as it is water-soluble. A mixture containing 10 mm of a model nucleophile, phenyl fosylate, and TMG was stirred in DMSO and the progress of the reactions was checked from GC-MS analysis. Among all nucleophiles tested in the experiments, only *p*-cresol was completely consumed and the desired product, *p*-tolyl phenyl sulfate, was isolated in 93.5% yield within 1.5 h (Entry 1). The same reaction in higher concentration (0.1 M) was complete in 5 min (data not shown). None of the corresponding products were generated for other model nucleophiles even in 12 h. Though 4-methylimidazole, a nucleophile corresponding to histidine (His), was converted to the corresponding sulfamide (Figure 5), adding a small amount of Ni^{2+}, which is an effective binder to imidazole, prevented the sulfamide formation almost completely. Through these several test reactions, we envisaged that the phenolic nucleophile of Tyr can chemoselectively participate in a modified SuFEx reaction among nucleophilic amino acids. Several SuFEx-mediated bioconjugations aided by binding affinity or UAA incorporation have been reported,[14, 23, 24] but Tyr selective reaction condition is unprecedented.

4.2. The SuFEx application in TAT 47-57 peptide labeling by fluorophore

We then assessed the possibility of the SuFEx application in peptide labeling. We took TAT 47–57, the most well-known cell penetrating peptide (CPP),[25-27] as a model substrate. TAT 47–57 possesses one Tyr at the N-terminus, which allowed us to attach a fluorescent tag using the SuFEx chemistry. We synthesized rhodamine-conjugated aryl fosylate (Rho-Fs) as a fluorescent tag (Figure 32). TAT 47–57, upon reaction with Rho-Fs, yielded Rho-TAT quantitatively. The detected *m/z*, 2183.19, by MALDI-TOF MS, matched well the expected mass of Rho-TAT (Figure 34). No multi-substituted products were generated (Figure 35). Though TAT peptide has several inherent Arg sidechains, *m/z* of Rho-TAT was not detected by MALDI-TOF in the absence of TMG even after 12 h-stirring (Figure 36). This result implies that, in a non-confined environment, addition of TMG was essential in this Tyr-selective SuFEx reaction.

To confirm that the SuFEx reaction took place selectively at Tyr in TAT 47–57, we compared MALDI-TOF/TOF spectrum of TAT 47–57 with that of Rho-TAT (Figure 37). The *N*-terminal fragment peaks of TAT 47–57 (b_8-b_{10}) were not found in the spectrum of Rho-TAT. Instead, different *N*-terminal fragment peaks (b_7-b_9) appeared in the Rho-TAT spectrum. The mass differences between 1) b_8 (1073.4) of TAT 47–57 and b_8 (1697.6) of Rho-TAT, and 2) b_9 (1229.6) of TAT 47–57 and b_9 (1853.9) of Rho-TAT were 624.2 and 624.3, respectively, which are corresponding to the calculated mass shift. From the analysis, we could validate the *N*-terminal Tyr is the conjugated site. Subsequently, we examined the cell permeability of Rho-TAT to *HeLa* cells. Rho-TAT exhibited efficient cell permeation and localization in the nucleus (Figure 38).[28-31] Accordingly, we concluded that the SuFEx chemistry would be a reasonable method for the selective chemical conjugation of two functional molecules (fluorescent molecule and CPP) in peptide chemistry without

affecting each other. Moreover, cell viability test was performed by synthesizing compound **12** to confirm the cytotoxicity of diaryl sulfate bond. As a result, it was confirmed that there was almost no cytotoxicity for 24 hours and 48 hours at a high concentration of 100 μM in the cells (Figure 44), and it is expected that this diaryl sulfate bond can be used as a suitable material *in vivo*.

4.3. The SuFEx application in PEGylation of rhEPO

After the successful functionalization of the short peptide, we proceeded further to protein functionalization in aqueous conditions. We selected recombinant human erythropoietin (rhEPO) as a model protein. EPO is a glycoprotein hormone, which induces red blood cell production. rhEPO is used for treating severe anemia caused by chronic kidney disease.[32] To improve the pharmacokinetic stability of rhEPO, the conjugation of poly(ethylene glycol) (PEG) is widely used.[33] For example, Mircera, one of the most well-known rhEPO PEGylated by the Lys-succinimide chemistry, shows a significantly prolonged half-life.[34] Although the succinimide predominantly reacts with either Lys-45 or Lys-52 among seven Lys residues on the surface of rhEPO, a heterogeneous mixture of isomers is produced. The intrinsic activity of each isomer has not been documented because of the purification difficulty. However, since only one Tyr (Tyr-49) exists on the surface of rhEPO, it was predicted that a modification of the residue was not related to the function or binding of rhEPO.[35] Therefore, Tyr-49 appeared to be a suitable site for the SuFEx-mediated PEGylation without any interruption in protein function (Figure 39).

We synthesized a short fosylated PEG (PEG-Fs; Mn=2,372, PD=1.004). rhEPO, PEG-Fs, and TMG were dissolved in Tris buffer (50 mM, pH 8.0) and the reaction mixture was incubated at ambient temperature. Adding excess TMG did not affect pH of the buffer solution at all. After 3 h, a MALDI-TOF MS spectrum showed a

clear shift of *m/z*, approximately 1980 Da (Figure 40). The shift corresponds to the molecular weight of the conjugated PEG, which supported the monomeric PEGylation on rhEPO. The PEGylation site was confirmed by an in-depth mass study after trypsin treatment. The data verified that the PEG was conjugated solely on Tyr-49 (Figure 42, Table 2 and Table 3) and none of the internal Tyr residues underwent PEGylation. This result implies that the structural integrity of EPO was maintained during the whole process. Interestingly, none of His was PEGylated under this condition without any additives such as Ni^{2+}, even though some exposed His residues (His-32 and His-94) are present on the protein surface.

Finally, we examined whether the hematogenous function of the PEGylated rhEPO (PEG-rhEPO) was retained after the PEGylation through SuFEx reaction. Native rhEPO and PEG-rhEPO were intravenously administered into Balb/c mice at 10 $mgkg^{-1}$ dose every three days. *In vivo* activities were compared by the measurement of the hematocrit (HCT).[36] The control group showed a significant decrease of the HCT level (ΔHCT=-2–5%) throughout 15 days. On the other hand, the PEG-rhEPO-treated group demonstrated slightly elevated level of the HCT during the period (Figure 43). The rhEPO-treated group showed a similar tendency of the HCT increase to that of the PEG-rhEPO-treated group. The statistical analysis showed clear difference in the HCT level between the control and the PEGrhEPO-treated groups and no significant difference was detected between PEG-rhEPO- and rhEPO-treated groups. The preserved *in vivo* activity of PEG-rhEPO strongly supports that the Tyr-specific modification through the SuFEx would be a novel and feasible strategy for the protein functionalization.

5. Conclusions

We have discovered a new chemoselective reaction for the Tyr functionalization in a model peptide and a protein. Among various nucleophilic amino acid residues, exclusive reaction of Tyr with aryl fosylate was observed. We clearly demonstrated the applicability of the SuFEx chemistry to the peptide and protein modifications without disrupting their own functions. Since the reaction condition of the SuFEx chemistry is simple, chemoselective, and non-destructive, we anticipate a broad utility of this Tyr-selective bioconjugation strategy in the field of chemical biology.

6. References

[1] P. C. Ng, S. Henikoff, *Nucleic Acids Res*. **2003**, *31*, 3812-3814.

[2] T. Proft, *Biotechnol. Lett*. **2010**, *32*, 1-10.

[3] K. Lang, J. W. Chin, *Chem. Rev*. **2014**, *114*, 4764-4806.

[4] N. Krall, F. P. d. Cruz, O. Boutureira, G. J. L. Bernandes, *Nat. Chem.* **2015**, *7*, 103-113.

[5] T. S. Young, I. Ahmad, J. A. Yin, P. G. Schultz, *J. Mol. Biol.* **2010**, *395*, 361-374.

[6] N. Jain, S. W. Smith, S. Ghone, B. Tomczuk, *Pharm. Res.* **2015**, *32*, 3526-3540.

[7] J. M. Chalker, G. J. Bernardes, Y. A. Lin, B. G. Davis, *Chem. Asian J.* **2009**, *4*, 630-640.

[8] N. E. Sharpless, M. Flavin, *Biochemistry.* **1966**, *5*, 2963-2971.

[9] J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi, M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, *45*, 5307–5311; *Angew. Chem.* **2006**, *118*, 5433-5437.

[10] L. S. Witus, C. Netirojjanakul, K. S. Palla, E. M. Muehl, C.-H. Weng, A. T. Iavarone, M. B. Francis, *J. Am. Chem. Soc.* **2013**, *135*, 17223-17229.

[11] S. Lin, X. Yang, S. Jia, A. M. Weeks, M. Hornsby, P. S. Lee, R. V. Nichiporuk, A. T. Iavarone, J. A. Wells, F. D. Toste, C. J. Chang, *Science* **2017**, *355*, 597-602.

[12] H. O. Villar, L. M. Kauvar, *FEBS Lett*. **1994**, *349*, 125-130.

[13] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed*. **2014**, *53*, 9430–9448; *Angew. Chem.* **2014**, *126*, 9584-9603.

[14] W. Chen, J. Dong, L. Plate, D. E. Mortenson, G. J. Brighty, S. Li, Y. Liu, A. Galmozzi, P. S. Lee, J. J. Hulce, B. F. Cravatt, E. Saez, E. T. Powers, I. A. Wilson, K. B. Sharpless, J. W. Kelly, *J. Am. Chem. Soc.* **2016**, *138*, 7353-7364.

[15] C.-E. Yeom, H. W. Kim, S. Y. Lee, B. M. Kim, *Synlett* **2007**, *1*, 0146-0150.

[16] H. Ban, J. Gavrilyuk, C. F. Barbas III, *J. Am. Chem. Soc.* **2010**, *132*, 1523-1525.

[17] N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942-15943.

[18] O. O. Fadeyi, L. R. Hoth, C. Choi, X. Feng, A. Gopalsamy, E. C. Hett, R. E. Kyne Jr, R. P. Robinson, L. H. Jones, *ACS Chem. Biol.* **2017**, *12*, 2015-2020.

[19] A. Younai, G. F. Chin, J. T. Shaw, *J. Org. Chem*. **2010**, *75*, 8333-8336.

[20] J. Gao, P. Wang, R. W. Giese, *Anal. Chem*. **2002**, *74*, 6397-6401.

[21] a) S. L. Pedersen, A. P. Tofteng, L. Malik, K.J. Jensen, *Chem. Soc. Rev*. **2012**, *41*, 1826-1844. b) G. S. Vanier, *Methods. Mol. Biol*. **2013**, *1047*, 235-249.

[22] a) G. Zhou, G. Luo, Y. Zhou, K. Zhou, X. Zhang, L. Huang, *Electrophoresis* **1998**, *19*, 2348-2355. b) E. Giménez, F. Benavente, C. de Bolós, E. Nicolás, J. Barbosa, V. Sanz-Nebot, *J. Chromatogr. A* **2009**, *1216*, 2574-2582.

[23] N. Wang, B. Yang, C. Fu, H. Zhu, F. Zheng, T. Kobayashi, J. Liu, S. Li, C. Ma, P. G. Wang, Q. Wang, L. Wang, *J. Am. Chem. Soc.* **2018**, *140*, 4995-4999.

[24] S. Li, D. Cohen-Karni, L. T. Beringer, C. Wu, E. Kallick, H. Edington, M. J. Passineau, S. averick, *Polymer* **2016**, *99*, 7-12.

[25] H. Brooks, B. Lebleu, E. Vives, *Adv. Drug Delivery Rev*. **2005**, *57*, 559-577.

[26] X. Zhang, Y. Jin, M. R. Plummer, S. Pooyan, S. Gunaseelan, P. J. Sinko, *Mol. Pharm*. **2009**, *6*, 836-848.

[27] D. M. Copolovici, K. Langel, E. Eriste, U. Langel, *ACS Nano* **2014**, *8*, 1972-1994.

[28] J. Y. Zhao, R. Cui, Z. L. Zhang, M. Zhang, Z. X. Xie, D. W. Pang, *Nanoscale* **2014**, *6*, 13126-13134.

[29] S. W. Jones, R. Christison, K. Bundell, C. J. Voyce, S. M. Brockbank, P. Newham, M. A. Lindsay, Br. *J. Pharmacol*. **2005**, *145*, 1093-1102.

[30] E. VivHs, P. Brodin, B. Lebleu, *J. Biol. Chem.* **1997**, *272*, 16010-16017.

[31] J. M. de la Fuente, C. C. Berry, *Bioconjugate Chem*. **2005**, *16*, 1176-1180.

[32] A. Sanchez-Fructuoso, L. Guirado, J. C. Ruiz, V. Torregrosa, E. Gonzalez, M. L. Suarez, R. Gallego, *Transplant. Proc.* **2010**, *42*, 2931-2934.

[33] Y. J. Wang, S. J. Hao, Y. D. Liu, T. Hu, G. F. Zhang, X. Zhang, Q. S. Qi, G. H.

Ma, Z. G. Su, *J. Controlled Release* **2010**, *145*, 306-313.

[34] I. C. Macdougall, K.-U. Eckardt, *Lancet* **2006**, *368*, 947-953.

[35] R. A. Cohan, A. Madadkar-Sobhani, H. Khanahmad, F. Roohvand, M. R. Aghasadeghi, M. H. Hedayati, Z. Barghi, M. S. Ardestani, D. N. Inanlou, D. Norouzian, *Int. J. Nanomed.* **2011**, *6*, 1217-1227.

[36] G. G. Kochendoerfer, S. Y. Chen, F. Mao, S. Cressman, S. Traviglia, H. Shao, C. L. Hunter, D. W. Low, E. N. Cagle, M. Carnevali, V. Gueriguian, P. J. Keogh, H. Porter, S. M. Stratton, M. C. Wiedeke, J. Wilken, J. Tang, J. J. Levy, L. P. Miranda, M. M. Crnogorac, S. Kalbag, P. Botti, J. Schindler-Horvat, L. Savatski, J. W. Adamson, A. Kung, S. B. H. Kent, J. A. Bradburne, *Science* **2003**, *299*, 884-887.

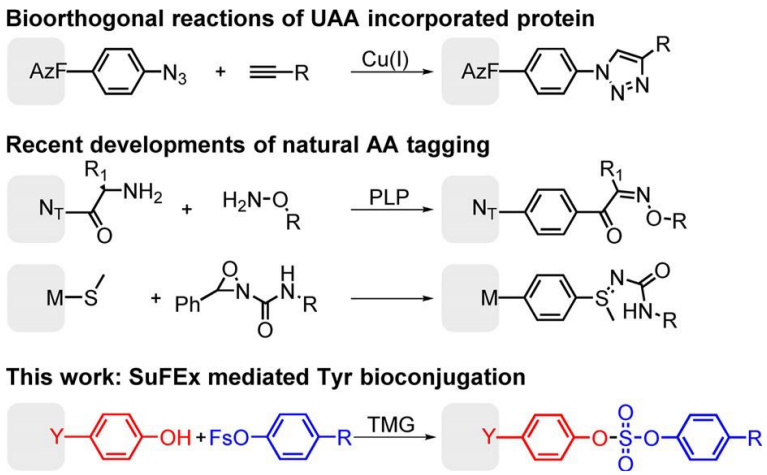


Figure 1. Artificial protein functionalization strategies. Protein functionalization using the click reaction of genetically modified proteins with an unnatural amino acid (first row) and recent discoveries of bioconjugations to natural amino acids in native proteins (second row). Selective Tyr modification in native proteins through the SuFEx chemistry.

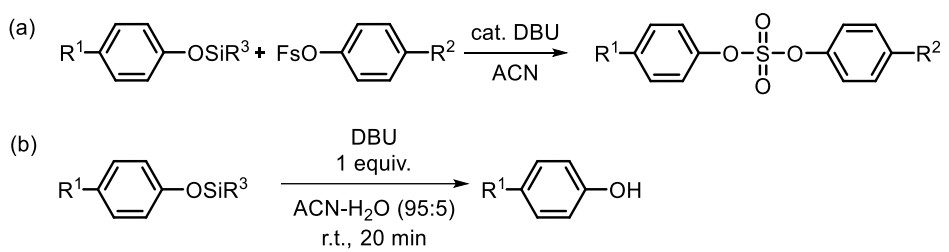


Figure 2. Comparison between SuFEx reaction developed by Sharpless group and stoichiometric DBU mediated desilylation reaction reported previously by us. (a) In SuFEx, the reaction between an aryl silyl ether and an aryl fosylate in the presence of a catalytic amount of DBU proceeds to sulfate click reaction. (b) Deprotection of aryl silyl ethers can be mediated through the use of a stoichiometric amount of DBU.

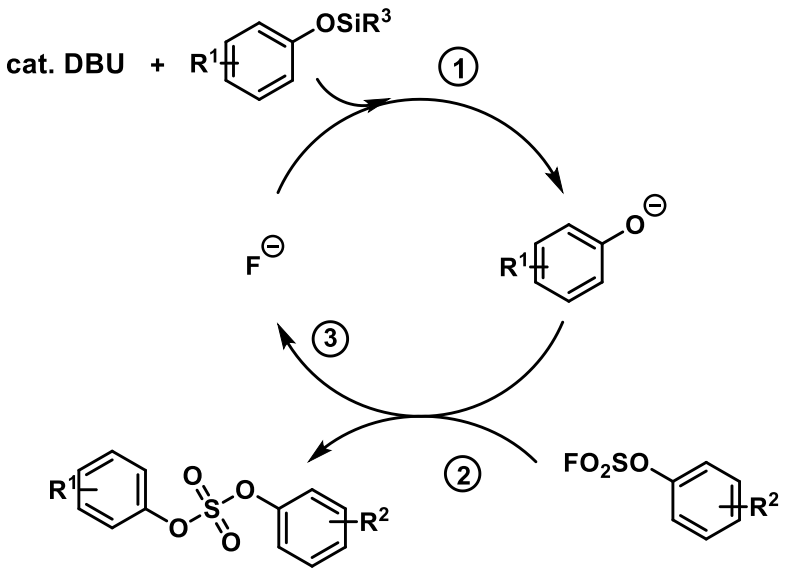


Figure 3. A proposed mechanism for SuFEx reaction. 1) Deprotection of an aryl silyl ether by DBU generating an aryloxy anion. 2) Formation of a sulfate product by the reaction between the aryloxy anion and an aryl fosylate and concomitant release of a fluoride anion. 3) The fluoride anion produced deprotects another aryl silyl ether to generate another aryloxy anion, which is ready to react with another fosylate, completing the catalytic cycle.

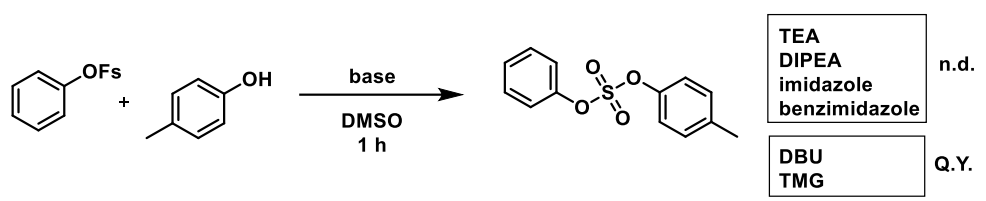


Figure 4. Base screening in modified SuFEx reaction between *p*-cresol and phenyl fosylate. Progress of the reaction was checked by GC-MS analysis.

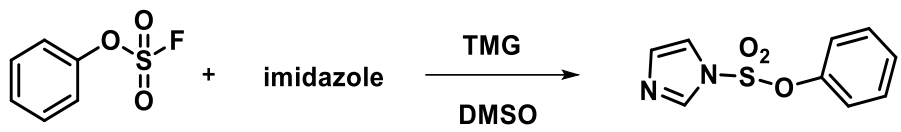


Figure 5. Reaction between phenyl fosylate and imidazole to yield a sulfamide product.

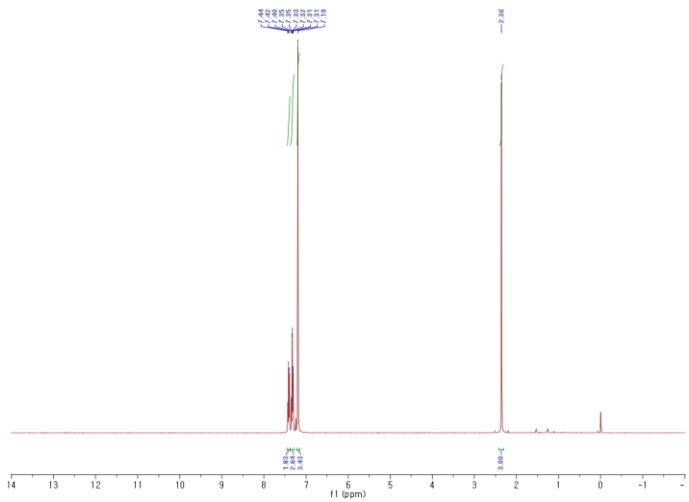


Figure 6. ^{1}H NMR spectrum of phenyl *p*-tolyl sulfate (**1**) in $CDCl_3$.

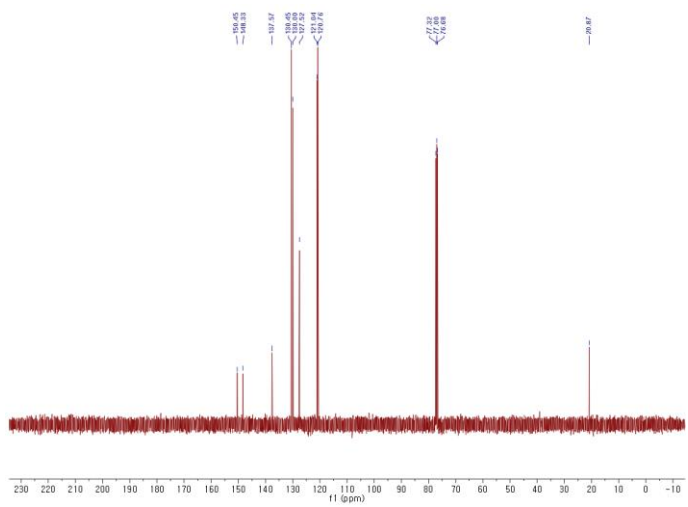


Figure 7. ^{13}C NMR spectrum of phenyl *p*-tolyl sulfate (1) in $CDCl_3$.

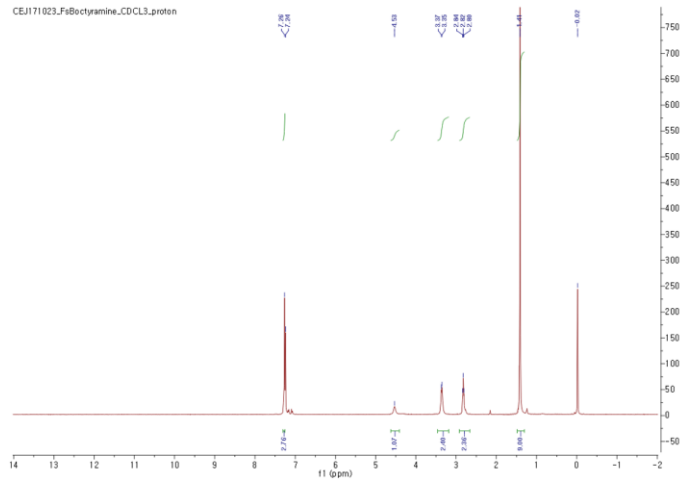


Figure 8. ^{1}H NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$.

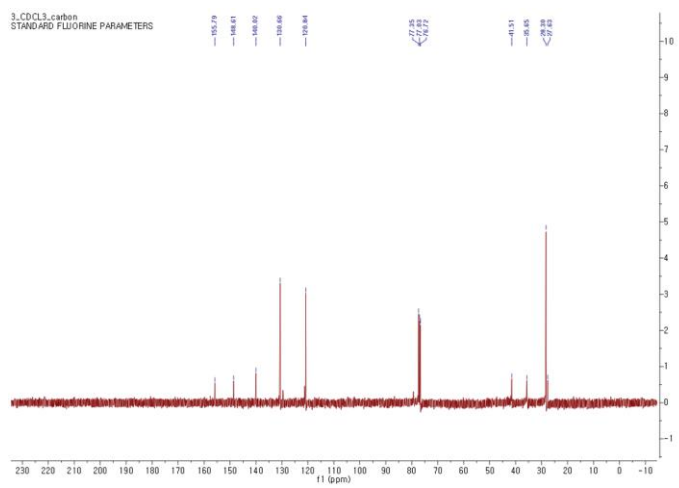


Figure 9. ^{13}C NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$.

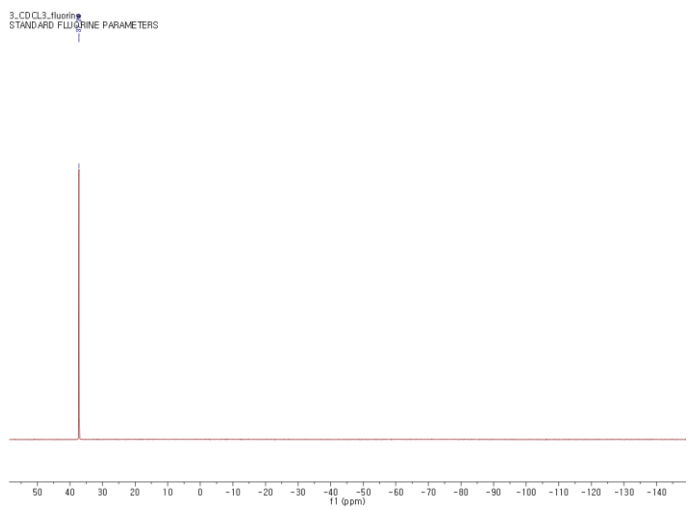


Figure 10. ^{19}F NMR spectrum of 4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (**3**) in $CDCl_3$.

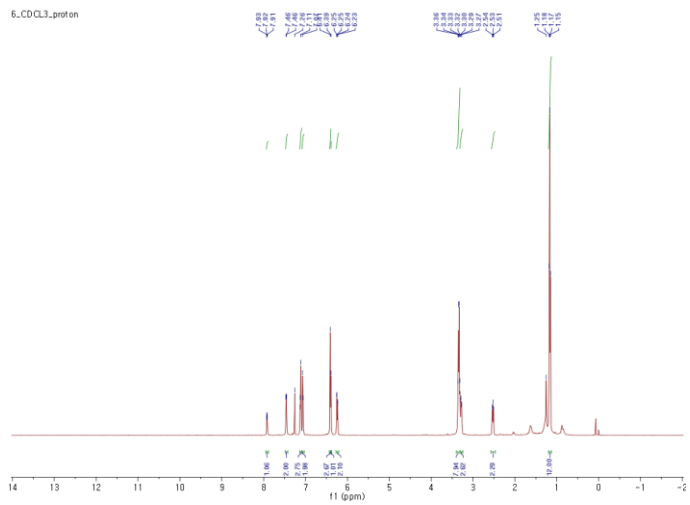


Figure 11. ^{1}H NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$.

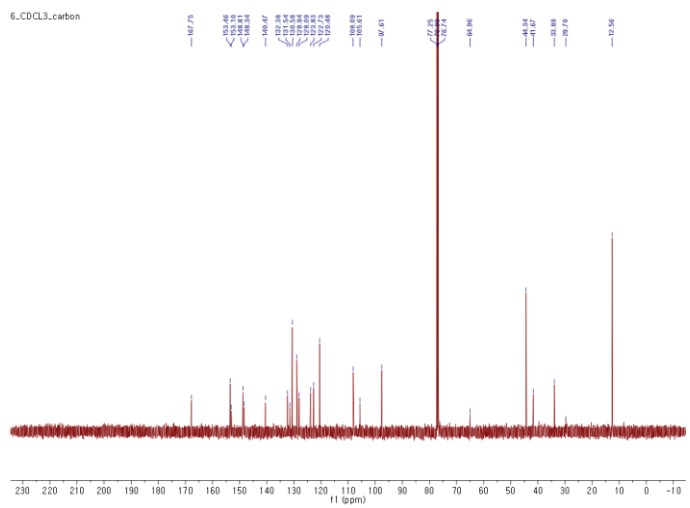


Figure 12. ^{13}C NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$.

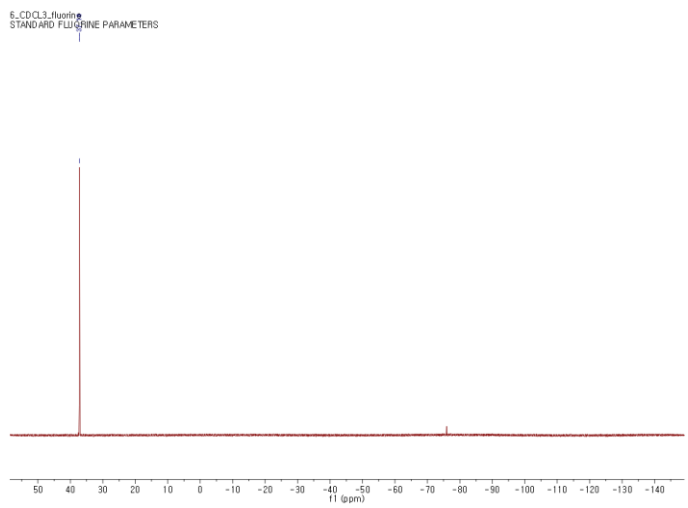


Figure 13. ^{19}F NMR spectrum of *N*-(6-(Diethylamino)-9-(2-((4-((fluorosulfonyl)oxy)phenethyl)carbamoyl)phenyl)-3H-xanthen-3-ylidene)-*N*-ethylethanaminium (**5**) in $CDCl_3$.

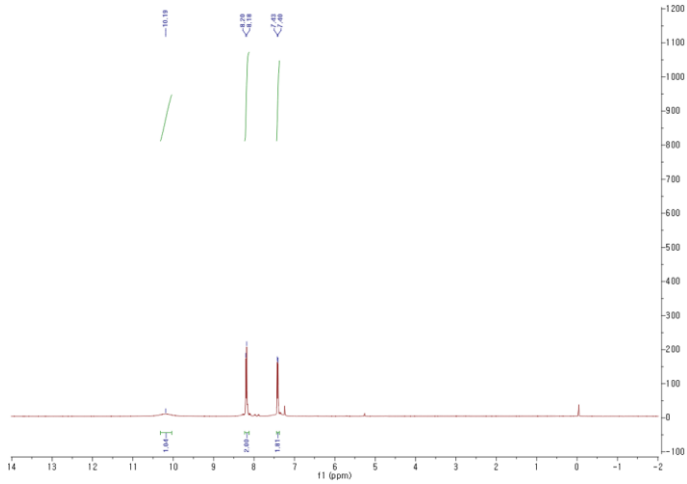


Figure 14. ^{1}H NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$.

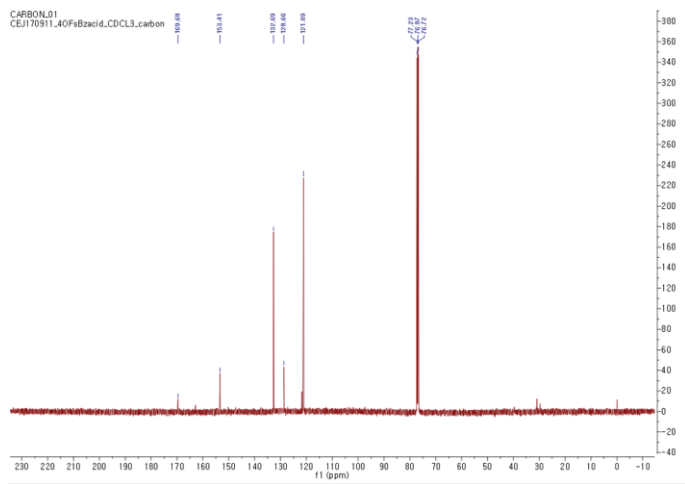


Figure 15. ^{13}C NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$.

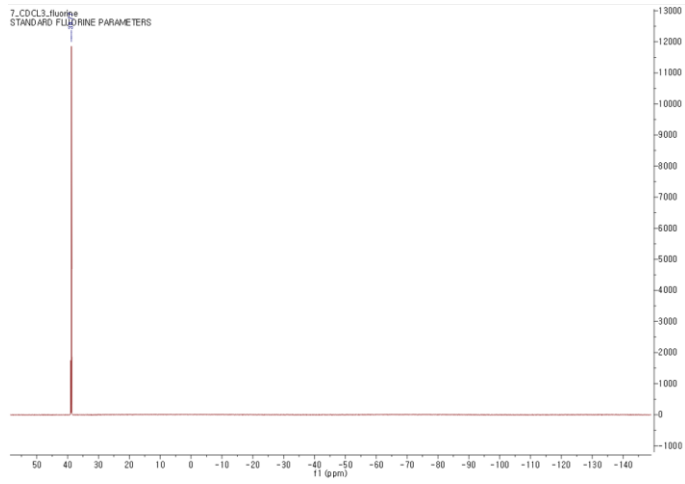


Figure 16. ^{19}F NMR spectrum of 4-((Fuorosulfonyl)oxy)benzoic acid (**6**) in $CDCl_3$.

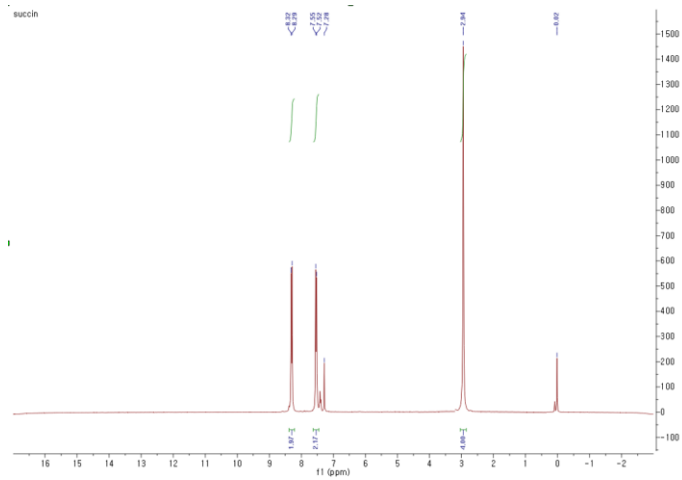


Figure 17. ^{1}H NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$.

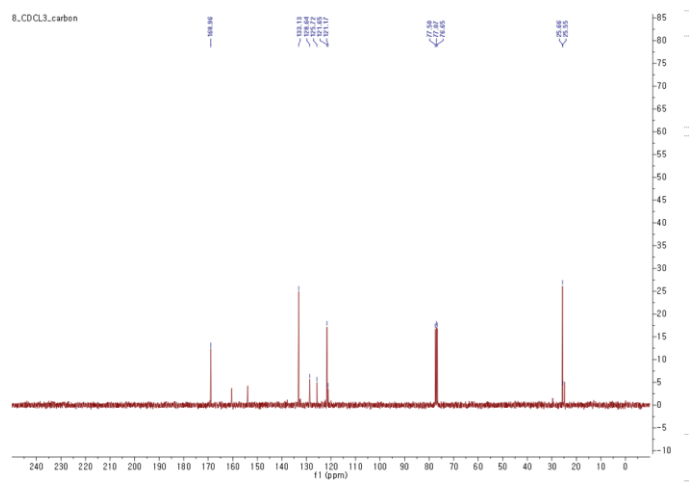


Figure 18. ^{13}C NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$.

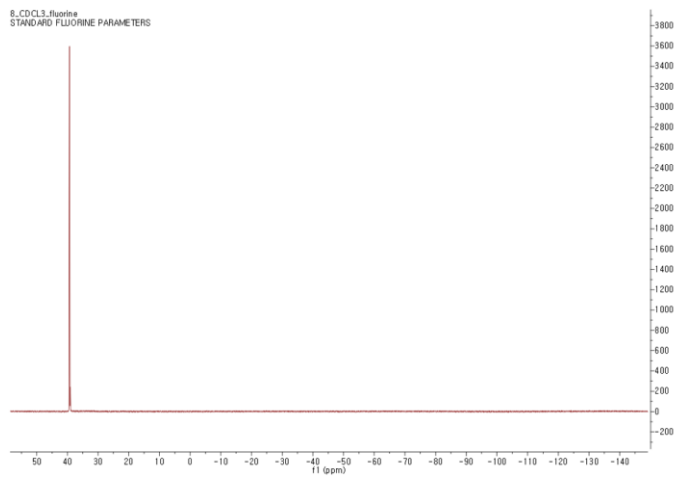


Figure 19. ^{19}F NMR spectrum of 2,5-Dioxopyrrolidin-1-yl-4-((fluorosulfonyl)oxy)benzoate (**7**) in $CDCl_3$.

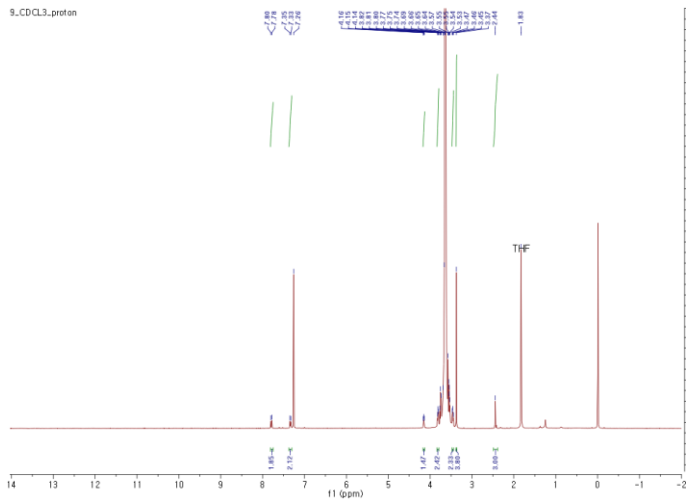


Figure 20. ^{1}H NMR spectrum of methoxy poly ethylene glycol tosylate (**8**) in $CDCl_3$.

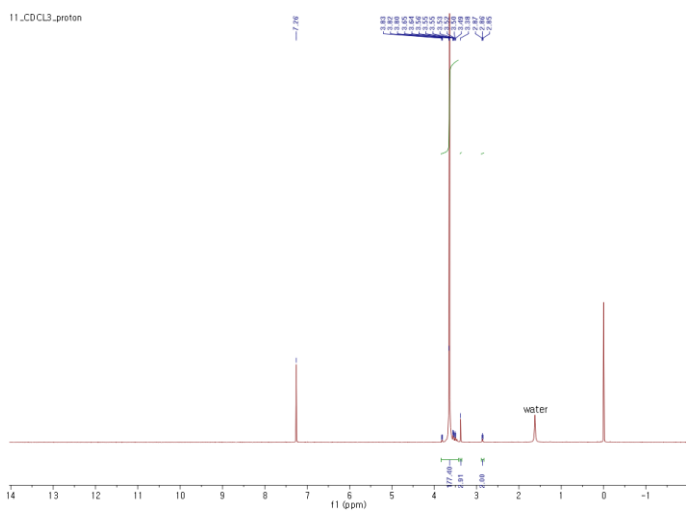


Figure 21. ^{1}H NMR spectrum of methoxy poly ethylene glycol amine (**10**) in $CDCl_3$.

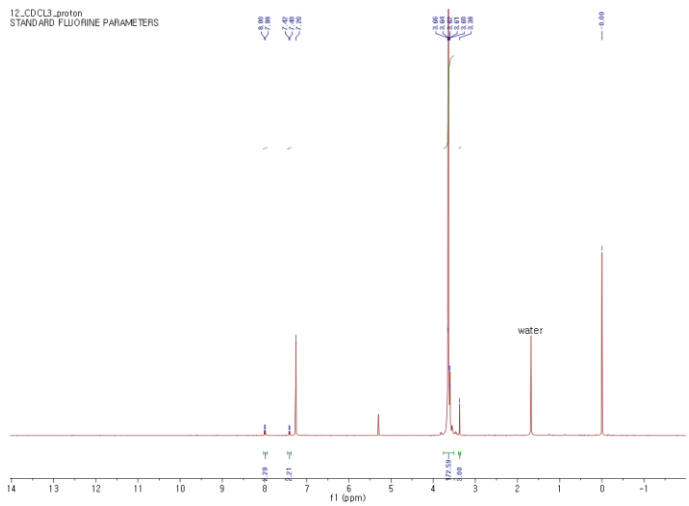


Figure 22. ^{1}H NMR spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (**11**) in $CDCl_3$.

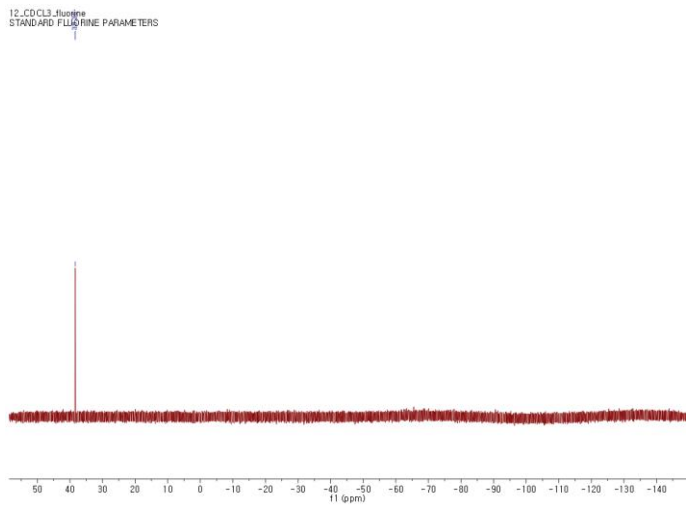


Figure 23. ^{19}F NMR spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (**11**) in $CDCl_3$.

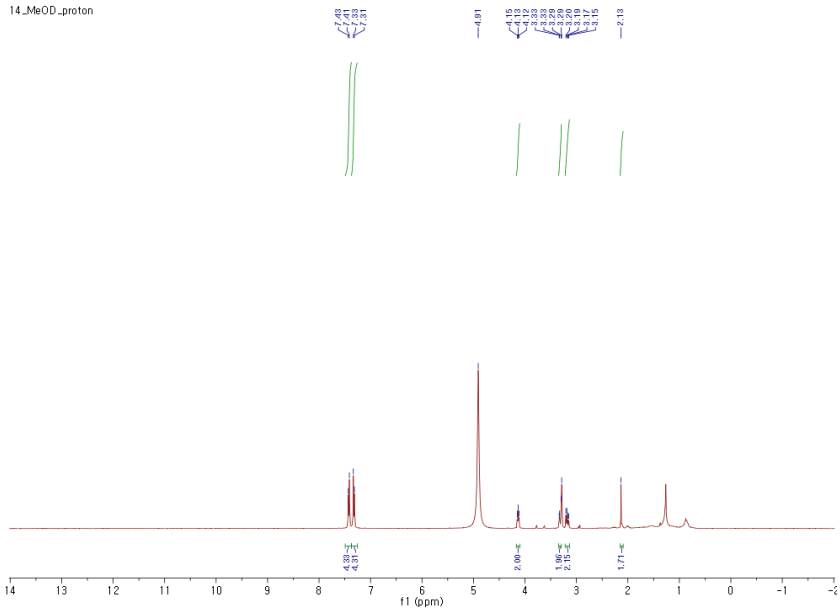


Figure 24. ^{1}H NMR spectrum of (2S,2'S)-3,3'-((Sulfonylbis(oxy))bis(4,1-phenylene))bis(2-aminopropanoic acid) (**12**) in CD_3OD.

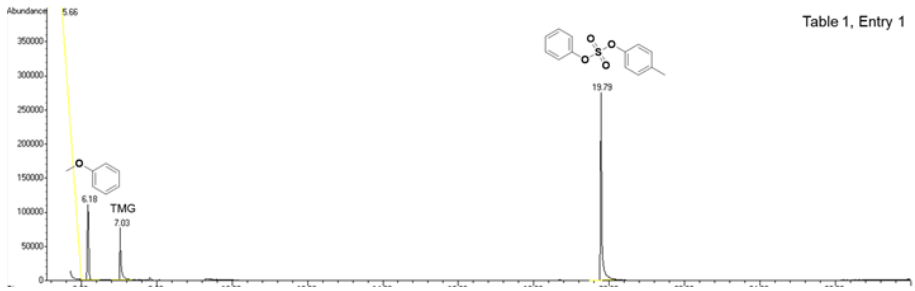


Figure 25. GC spectrum of entry 1 in table 1.

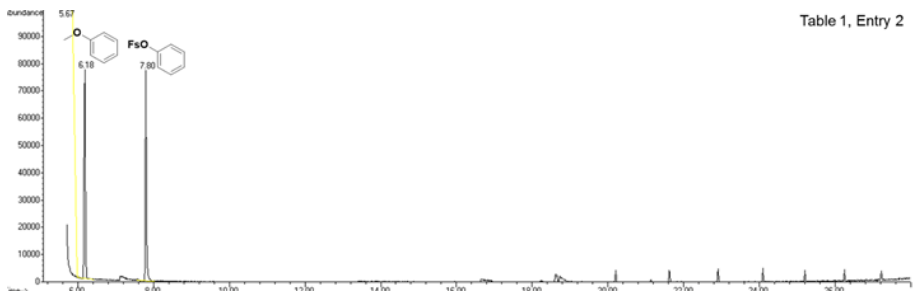


Figure 26. GC spectrum of entry 2 in table 1.

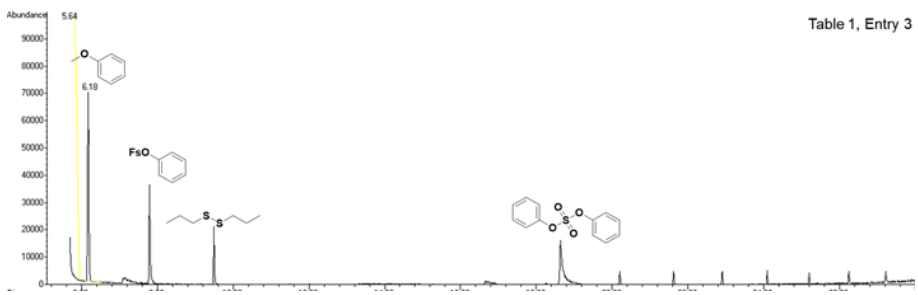


Figure 27. GC spectrum of entry 3 in table 1.

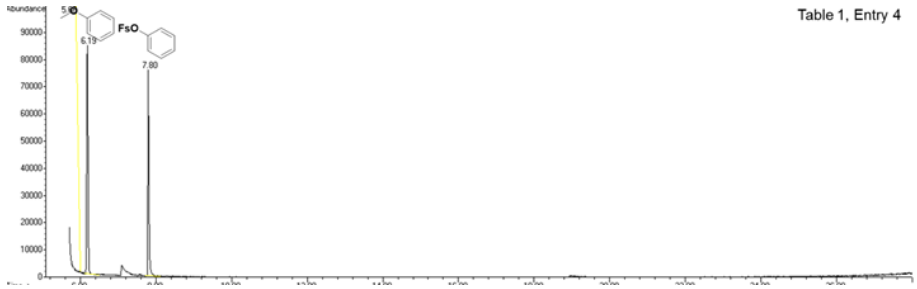


Figure 28. GC spectrum of entry 4 in table 1.

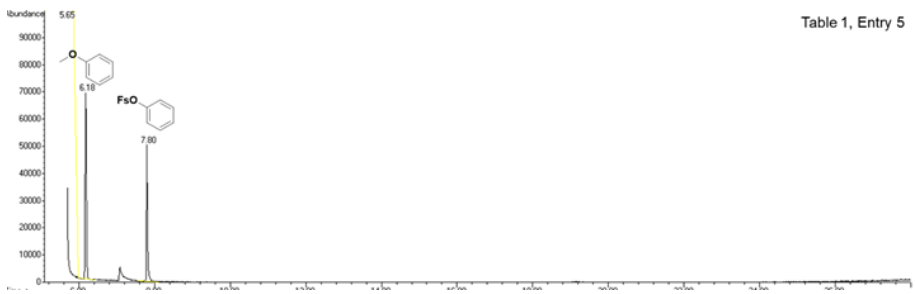


Figure 29. GC spectrum of entry 5 in table 1.

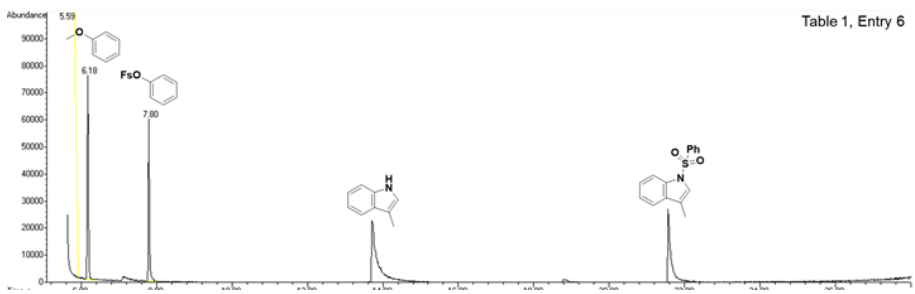


Figure 30. GC spectrum of entry 6 in table 1.

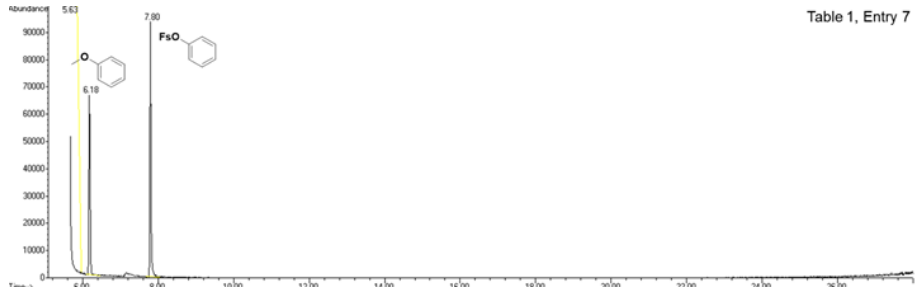


Figure 31. GC spectrum of entry 7 in table 1.

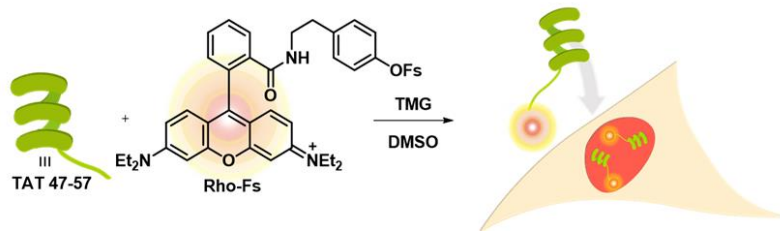


Figure 32. Functionalization of TAT 47–57 with a fluorescent small molecule.

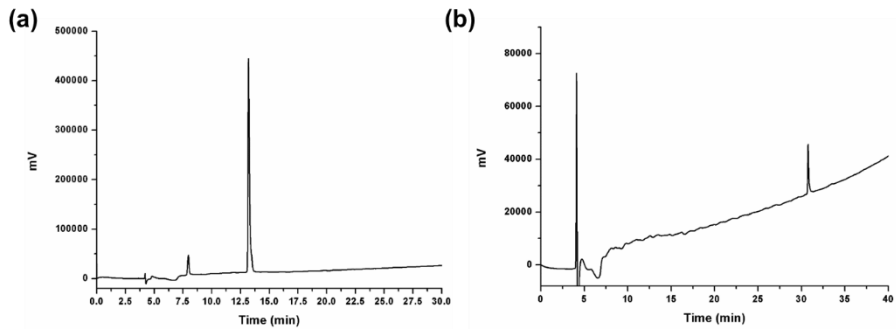


Figure 33. The HPLC chromatogram of TAT peptide (a) and Rho-TAT peptide (b). Purity of TAT peptide is 93% and that of Rho-TAT is 99%. Absorbance was measured at 220 nm.

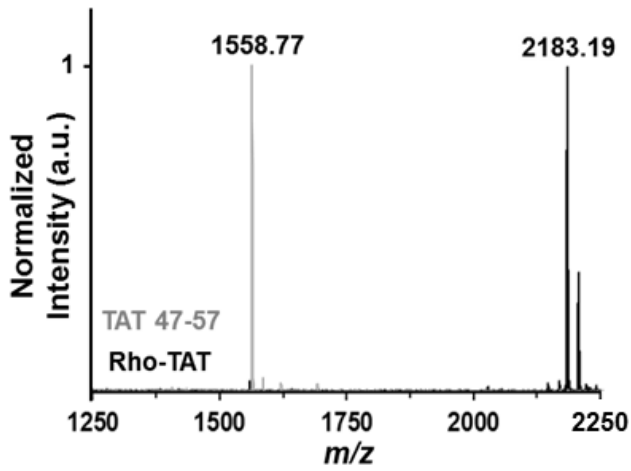


Figure 34. MALDI-TOF MS spectra of TAT 47–57 (*m/z* expected: 1558.97/observed: 1558.77) and Rho-TAT (expected: 2183.23/observed: 2183.19).

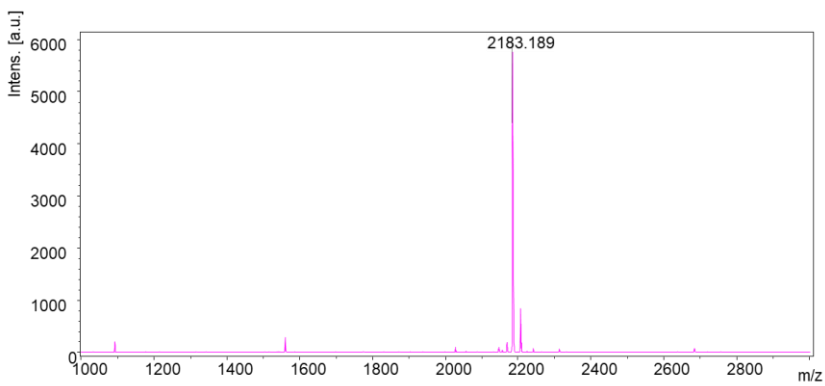


Figure 35. Large range MALDI-TOF MS spectrum of Rho-TAT peptide.

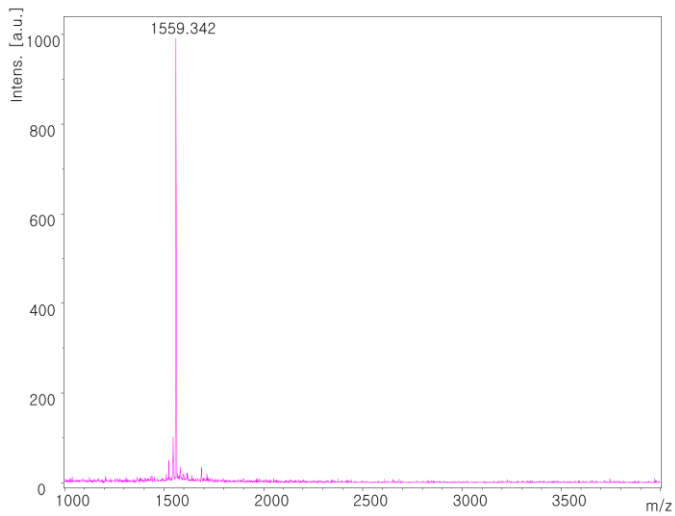


Figure 36. MALDI-TOF MS spectra of reaction mixture of TAT 47-57 and Rho-Fs without TMG.

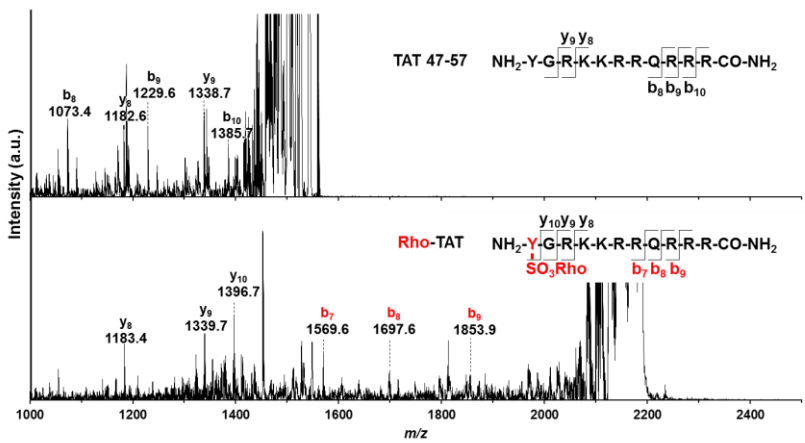


Figure 37. MALDI-TOF/TOF spectra of TAT 47-57 (upper) and Rho-TAT.

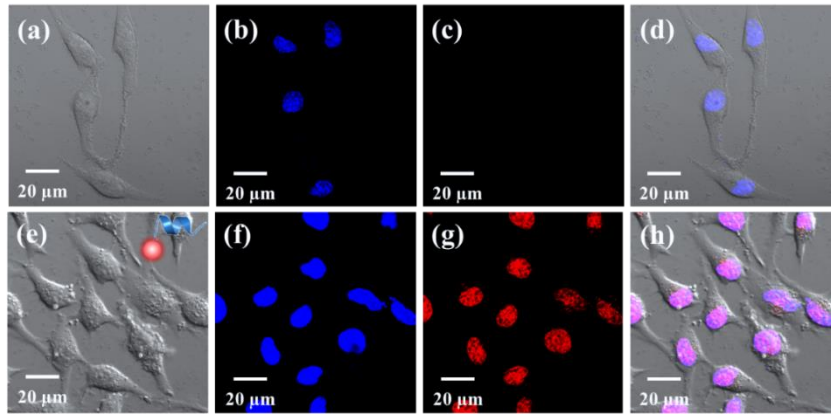


Figure 38. CLSM images of *HeLa* cells after 12 h-treatment of deionized water (a-d), and Rho-TAT peptide (e-h). Nucleus was stained with Hoechst 33342 (blue).

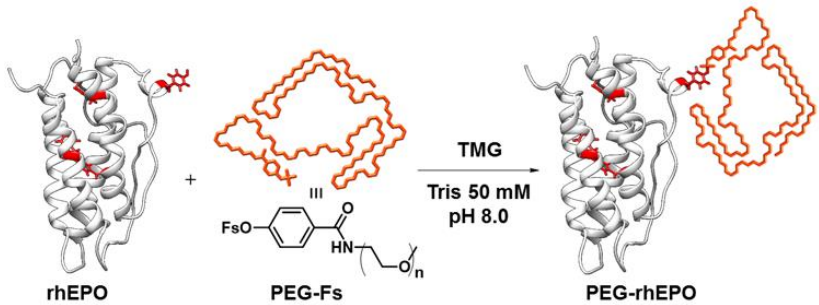


Figure 39. The schematic representation of the SuFEx reaction between rhEPO and a PEGylating reagent.

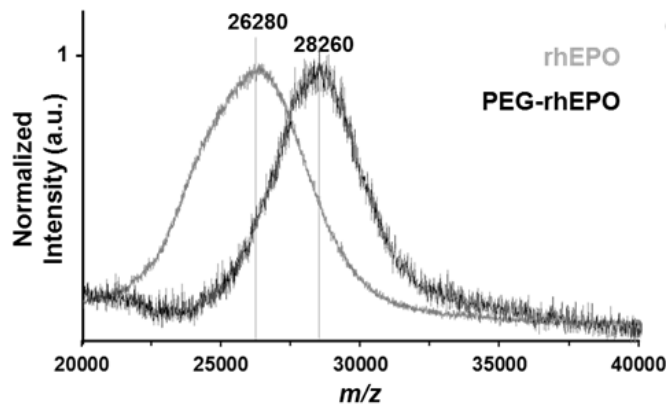


Figure 40. MALDI-TOF MS spectra of rhEPO (*m/z* observed; 26,280) and PEG-rhEPO (observed; 28,260).

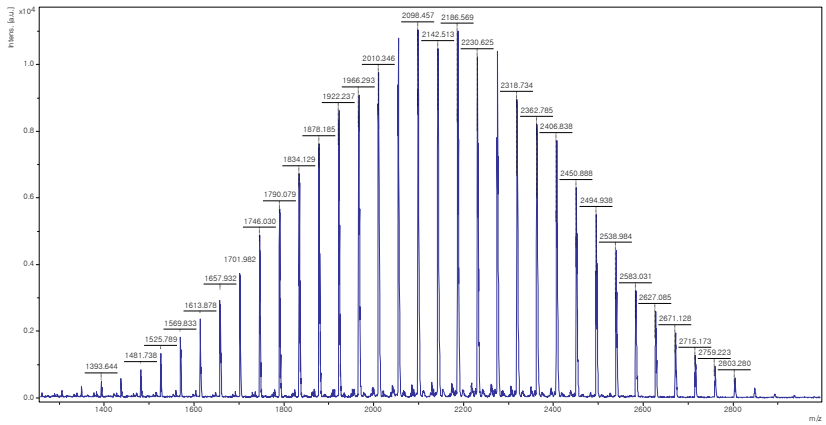


Figure 41. MALDI-TOF MS spectrum of 4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (11)

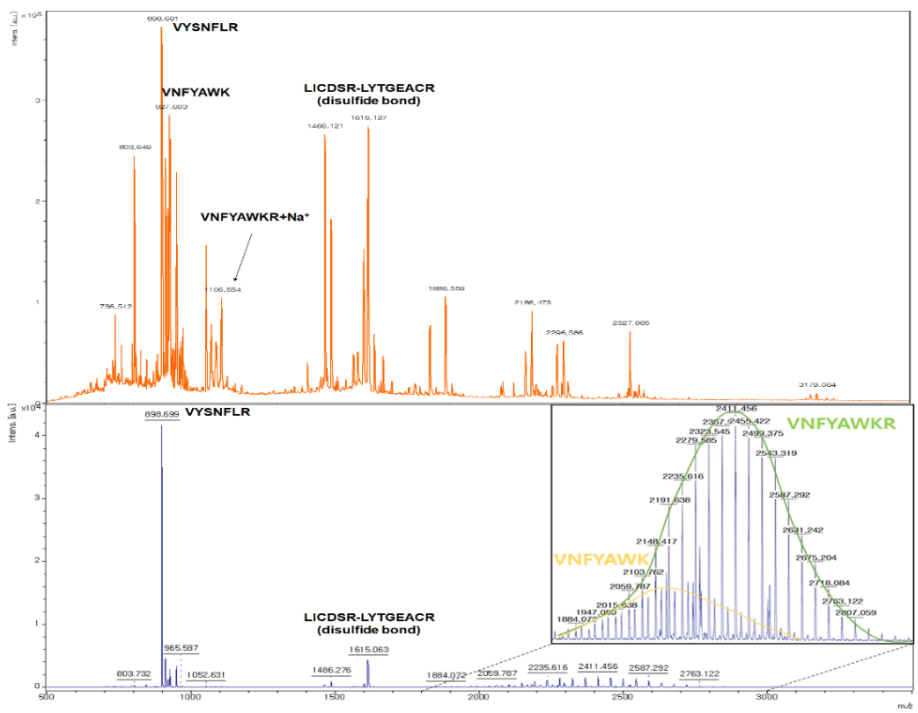


Figure 42. MALDI-TOF MS spectra of rhEPO (upper) and PEG-rhEPO after trypsin treatment.

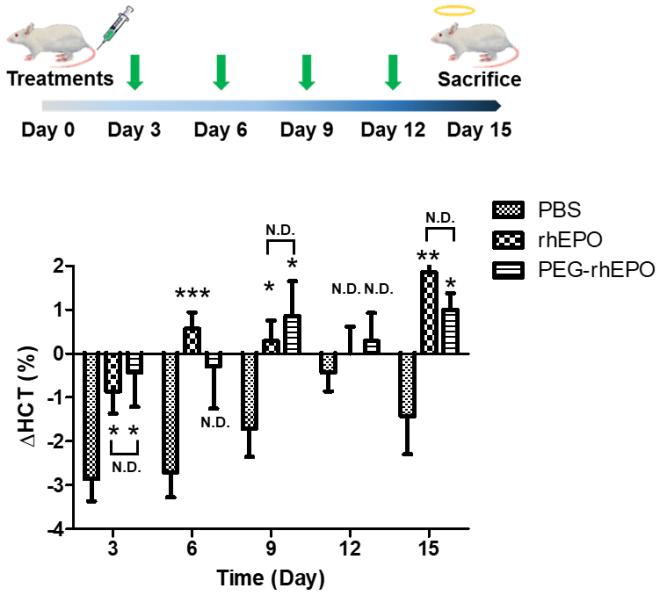


Figure 43. The HCT profiles of control, rhEPO-treated, PEG-rhEPO-treated Balb/c mice.

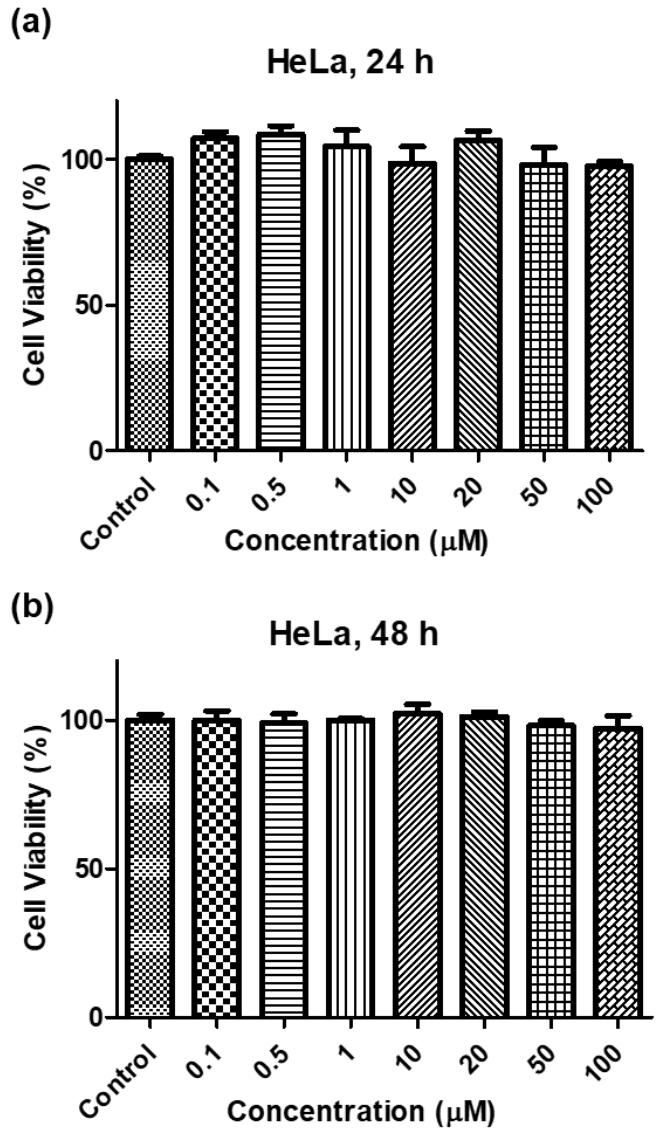
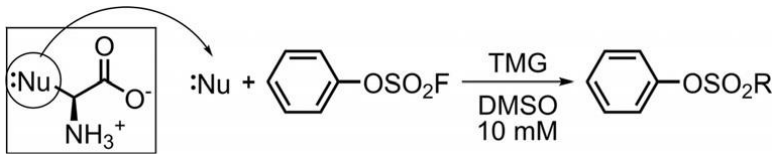


Figure 44. Dose dependent cytotoxicity of **12** in *HeLa* cells for 24 h (a) and 48 h (b).

Table 1. Comparison of SuFEx reactivity in the reactions of various model nucleophiles representing amino acids with phenyl fosylate.



Entry	Model nucleophile	Amino acid	Time (h)	Yield[a] (%)
1	*p*-cresol	Y	1.5	93.5
2	*n*-butylamine	K	12	n.d.
3	Propanethiol	C	12	n.d.
4	Methanol	S	12	n.d.
5	*N*-propylguanidine	R	12	n.d.
6	3-methylindole	W	12	12.5
7[b]	4-methylimidazole	H	12	trace

[a] Yield of isolated product. [b] 0.25 equiv of $NiCl_2(H_2O)_6$ was added.

Table 2. PEGylated VNFYAWKR *m/z* peaks in MALDI-TOF spectrum of PEG-rhEPO after trypsin treatment.

n (# of PEG)	15	16	17	18	19	20	21	22	23	24
Expected	1970.991	2015.018	2059.044	2103.070	2147.096	2191.122	2235.149	2279.175	2323.201	2367.227
Observed	1971.885	2015.838	2059.787	2103.762	2146.624	2191.638	2235.616	2279.858	2323.545	2367.508
n (# of PEG)	25	26	27	28	29	30	31	32	33	34
Expected	2411.254	2455.280	2499.306	2543.332	2587.358	2631.385	2675.411	2719.437	2763.463	2807.489
Observed	2411.456	2455.422	2499.375	2543.319	2587.292	2631.242	2675.204	2718.084	2763.122	2807.059

Table 3. PEGylated VNFYAWK *m/z* peaks in MALDI-TOF spectrum of PEG-rhEPO after trypsin treatment.

n (# of PEG)	18	19	20	21	22	23	24
Expected	1946.969	1990.995	2035.021	2079.048	2123.074	2167.100	2211.126
Observed	1947.050	1990.996	2034.946	2078.901	2122.854	2166.805	2211.492
n (# of PEG)	25	26	27	28	29	30	31
Expected	2255.153	2299.179	2343.205	2387.231	2431.257	2475.284	2519.31
Observed	2254.748	2298.679	2342.610	2386.546	2430.476	2474.422	2518.322

List of Publications

1. **Jung, D.**; Kim, Y.-J.; Lee, J.-K., Novel Strategy for Maintenance of Catalytic Activity Using Wrinkled Silica Nanoparticle Support in Fischer-Tropsch Synthesis. *Bulletin of the Korean Chemical Society* **2016,** *37* (3), 386-389.

2. Lim, J.; Pyo, J.; **Jung, D.**; Jung, H.-S.; Lee, J.-K., Preparation of mono-dispersed spherical titania nanoparticles with precise size control using ethylene glycol. *Journal of Sol-Gel Science and Technology* **2016,** *79* (1), 89-97.

3. Kang, J. S.; Lim, J.; Rho, W. Y.; Kim, J.; Moon, D. S.; Jeong, J.; **Jung, D.**; Choi, J. W.; Lee, J. K.; Sung, Y. E., Wrinkled silica/titania nanoparticles with tunable interwrinkle distances for efficient utilization of photons in dye-sensitized solar cells. *Sci. Rep.* **2016,** *6*, 30829.

4. Song, Y.; **Jung, D.**; Kang, S.; Lee, Y., Amine-selective affinity resins based on pH-sensitive reversible formation of covalent bonds. *Soft Matter.* **2017,** *13* (12), 2295-2298.

5. Jin, Y. J.; Kang, S.; Park, P.; Choi, D.; Kim, D. W.; **Jung, D.**; Koh, J.; Jeon, J.; Lee, M.; Ham, J.; Seo, J. H.; Jin, H. R.; Lee, Y., Anti-inflammatory and Antibacterial Effects of Covalently Attached Biomembrane-Mimic Polymer Grafts on Gore-Tex Implants. *ACS Appl. Mater. Interfaces* **2017,** *9* (22), 19161-19175.

6. Choi, E. J.; **Jung, D.**; Kim, J. S.; Lee, Y.; Kim, B. M., Chemoselective Tyrosine Bioconjugation through Sulfate Click Reaction. *Chem. Eur. J.* **2018,** *24* (43), 10948-10952.

7. Kang, M.; Kim, S.; Kim, H.; Song, Y.; **Jung, D.**; Kang, S.; Seo, J. H.; Nam, S.; Lee, Y., Calcium-Binding Polymer-Coated Poly(lactide- co-glycolide) Microparticles for Sustained Release of Quorum Sensing Inhibitors to Prevent Biofilm Formation on Hydroxyapatite Surfaces. *ACS Appl. Mater. Interfaces* **2019,** *11* (8), 7686-7694.

8. **Jung, D.**; Choi, D.; Sim, C.; Kim, Y.; Kang, S.; Nam, S. H.; Jang, J.; Kim, D.; Chang, M. S.; Park, J. U.; Lee, Y., De novo formation of citrate-based fluorophores on N-termini of peptides and proteins in cells and tissues. *Chem. Commun. (Camb)* **2020,** *56* (1), 74-77.

Abstract (국문 초록)

기능을 갖는 분자를 이용하여 생체분자에 표시하는 것은 복잡한 생명현상을 이해하고 설명하는데 중요한 역할을 하고 바이오이미징과 생물치료학의 발전을 가져왔다. 특히, 기능성 분자 중 형광체는 펩타이드 및 단백질에 표시하여 바이오이미징 분야에서 세포 내 생체분자의 거동을 관찰을 가능하게 하였고 비외과적인 방법으로 질병의 위치와 경과를 실시간으로 확인 할 수 있게 하였다. 또한 생물치료학 분야에서 항암제와 같은 치료 효과를 갖는 기능성 분자는 항체와 결합시켜 암세포만 선택적으로 치료를 할 수 있으며 폴리에틸렌글리콜과 같은 기능성 고분자는 단백질에 결합하여 구조를 안정화하고 생물학적 활성을 향상시켜준다.

이러한 기능성 분자를 생체분자에 결합시키는 연구에 많이 사용되는 방법은 생체분자의 천연 아미노산과 선택적인 생체접합을 통해 두 물질이 안정하게 결합하는 것과 유전암호를 통해 비천연 아미노산을 특정 자리에 도입하여 결합하는 것이 있다. 천연 아미노산에 생체접합을 하기 위한 방법은 아미노산 곁사슬을 화학적인 표적으로 하여 결합을 형성하는 것이다. 특히, 라이신 아미노산의 아민 그룹과 친핵 반응을 통한 결합을 많이 이용한다. 라이신의 아민과 선택적으로 결합하기 위해서는 결합 대상의 물질에 *N*-하이드록시숙신이미드, 이소티오시아네이트, 이소시아네이트 그리고 이미도에스터 등의 다양한 화학적 작용기를 도입하여 아민과 반응시킨다. 그밖에 시스테인의 싸이올과 말레이미드의 반응을 통해 결합시키는 방법을 많이 사용되고 있다. 하지만 대부분의 단백질에는 많은 라이신이 표면에 존재하기 때문에 반응 후에 활성이 다른 다양한 생성물이 만들어지는 문제점이 있고 시스테인의 경우 쉽게 산화되어 대부분이 디설파이드 결합 형태로 단백질에 존재하기때문에 반응 전에 환원 과정이 필요하다. 따라서, 생체분자와 기능성 분자를 선택적으로 접합을 할 수 있는 방법을 개발하는 것은 의미가 있다.

본 박사학위 논문은 1) 펩타이드와 단백질의 *N*-말단에 시트르산 염을 기반한 형광체를 드 노보로 형성함과 2) 황산염 클릭 반응을 통한 타이로신의 선택적인 생체접합 개발을 포함한다.

기존에 알려진 5-옥소-2,3-디하이드로-5H-[1,3]티아졸로[3,2-*a*]피리딘-3,7-디카복실산 형광체를 형성하는 방법은 고온에서

시트르산염과 시스테인의 탈수 반응과 축합 반응이다. 이 반응 과정에서 시트르산염과 시스테인은 큰 에너지가 필요한 아마이드 결합을 하게 된다. 이러한 이유로 아마이드 결합 형성을 촉진시켜주는 커플링 시약을 넣어줌으로써 펩타이드 합성조건에서 마이크로웨이브 방법과 상온에서도 형광체가 형성됨을 확인하였다. 이러한 결과로 *N*-말단에 시스테인을 도입한 펩타이드를 시트르산염과 커플링 시약을 사용하여 온화한 조건에서 드 노보로 형광체를 형성됨을 확인하였다. 더 나아가 커플링 시약을 넣어주는 반응 조건으로 고정된 세포와 조직 안에 존재하는 단백질과 펩타이드의 *N*-말단 아미노산들과 시트르산염이 반응하여 형광체가 형성 됨을 공초점 레이저 주사 현미경으로 확인하였다.

황산염 클릭 반응은 6가의 황과 불소의 교환을 통해 진행된다. 기존의 황산염 클릭 반응은 방향족의 플루오르황산염과 아릴 실릴 에테르 사이의 반응으로 알려져 있지만 염기를 이용하여 실릴 그룹 없이 직접적으로 방향족 수산화 음이온으로 활성화시켜 황산염 클릭 반응을 한 사례가 없었다. 따라서 본 연구에서는 다양한 아미노산 중에 타이로신에만 선택적으로 황산염 클릭 반응이 되도록 반응 조건을 개발하였다. 이러한 반응 조건으로 TAT 47-57 세포 투과성 펩타이드에 존재하는 타이로신에만 형광체를 선택적으로 생체접합을 할 수 있었고 세포에 형광-펩타이드를 처리한 뒤 공초점 레이저 주사 현미경으로 세포 내 핵에서 형광을 확인하였다. 더 나아가 빈혈 치료제로 알려진 에리트로포이에틴 단백질 표면에 노출된 49번 타이로신에만 폴리에틸렌글리콜 고분자를 황산염 클릭 반응으로 생체접합을 하였다. 동물실험을 통해서 고분자가 접합된 단백질을 생체 내로 투여 했을 때 혈액에서 적혈구가 차지하는 용적의 비중이 증가함을 확인하였다.

본 연구에서 새로운 생체유기화학 전략으로 펩타이드 및 단백질 내에 선택적으로 기능성 분자 구조를 형성하는 것은 바이오이미징과 생물치료학 분야에 기여를 할뿐만 아니라 화학 생물학 영역에도 광범위하게 사용될 것으로 예상한다.

주요어: 생체접합, 형광체, *N*-말단, 드 노보, 커플링 시약, 황산염 클릭 반응, 타이로신
학번: 2013-30907