이학박사학위논문

탄소-수소 활성화 및 설페이트 클릭 반응의 화학생물학 응용

Application of C-H Activation and Sulfate Click Reaction to Chemical Biology

2018 년 2 월

서울대학교 대학원

화학과 유기화학전공

최 은 정

Abstract

Application of C-H Activation and Sulfate Click Reaction to Chemical Biology

Choi, Eun Joung
Department of Chemistry, Organic Chemistry
The Graduate School
Seoul National University

Factors affecting cellular phenomena comprise a series of complicated and elaborate signal transductions. Due to the complexity of biomolecules, direct observation of a specific biomolecule is extremely difficult. Therefore conjugation of functionality onto the biomolecule is often implemented for the purpose of monitoring a bio-target. Since hydrophobic interaction between a biomolecule and a ligand is usually not strong enough to be monitored by common analytical tools, covalent bond formation of an artificial tag molecule has been developed.

After Huisgen's discovery of 1,3-dipolar cycloaddition between an azide and an alkyne, Sharpless group reported that the reaction can be catalyzed by a Cu(I) species. Furthermore, introduction of azidophenylalanine into proteins through expanded genetic code renders possible the conjugation of an artificial functionality such as fluorescence, and anti-cancer activity, etc. It is of particular note that the fluorescence attachment makes the spatiotemporal observation of bio-target possible. As a consequence, both the development of a new bioconjugation method, and the synthesis of functional molecules to be

ligated have become significant research subjects in chemical biology field.

This dissertation includes research results of two parts, 1) developing synthetic method of fluorescent molecules which can be used as a biosensor, and 2) discovery of novel tyrosine bioconjugation chemistry.

The first part shows one of our approaches in utilizing novel metal assisted 1,3-cycloaddition and Pd-mediated C-H activation for efficient syntheses of indolizinone fluorophores containing a small molecule library. Since the photophysical property of indolizinone is regulated by the electronic characteristics of C7 and C9 substituents, diversification of those substituents was crucial for the construction of a fluorescent molecule library. Through the new synthetic route, synthesis of a number of fluorescent molecules was accomplished, as diversification of C9 substituent occurs at the late stage of the synthesis. Moreover, several biosensors were synthesized, which had been previously infeasible.

In part 2, tyrosine selective conjugation method through sulfate click reaction is described. Since SuFEx (sulfur(VI) fluoride exchange) chemistry was known as the reaction between aryl fluorosulfate and aryl silyl ether, aryloxy anion has scarecely been recognized as a useful reaction partner for SuFEx. In this part, a reaction condition to allow for the reaction of aryl fluorosulfate with tyrosine in the presence of several nucleophilic amino acid residues was carefully selected, and a feasibility test of a new site-specific biomolecule modification was successfully carried out on a model peptide, TAT 47-57, and a physiologically important protein, erythropoietin.

keywords: Fluorescence, C-H activation, Bioconjugation, Tyrosine, SuFEx.
Student Number: 2011-20308.

Table of Contents

Abbreviations

TEA: trimethylamine
DIPEA: diisopropylethylamine
DCM: dichloromethane
MC: methylene chloride
EA: ethyl acetate
HEX: hexane
DMF: *N,N*-dimethylformamide
AcOH: acetic acid
TFA: trifluoro acetic acid
DBU: 1,8-Diazabicyclo(5.4.0)undec-7-ene
DDQ: 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
Boc: *t*-butyloxycarboxyl
TMS: trimethylsilyl
TBS: *t*-butyldimethylsilyl
BPin: pinacolato boryl
NMR: nuclear magnetic resonance
EPR: electron paramagnetic resonance
PeT: photoinduced electron transfer
PBS: phosphate-buffered saline
ESI: electron spray ionization
LRMS: low resolution mass spectroscopy
HRMS: high resolution mass spectroscopy
SuFEx: sulfur(VI) fluoride exchange
HOMO: highest occupied molecular orbital
LUMO: lowest unoccupied molecular orbital
UAA: unnatural amino acid
Tyr: tyrosine

Cys: cysteine
Met: methionine
His: histidine
Lys: lysine
Ser: serine
Thr: threonine
Trp: Tryptophan
CPP: cell penetrating peptide
rhEPO: recombinant human erythropoietin
Q.Y.: quantitative yield
n.d.: non detected
TMG: 1,1,3,3-Tetramethylguanidine
DPBS: Dulbecco's phosphate-buffered saline
FBS: fetel bovine serum
DMEM: Dulbecco's modified eagle medium
CLSM: confocal laser scanning microscopy
GC/MS: gas chromatography mass spectroscopy
LC/MS: liquid chromatography mass spectroscopy
MALDI: matrix assisted laser desorption/ionization
TOF: time of flight

List of Tables

List of Schemes

Scheme 3.5. Proposed selective Tyr modification in native proteins through the SuFEx chemistry.

Scheme 3.6. Reaction between phenyl fosylate and imidazole.

List of Figures

Chapter 1. Introduction

1.1 Bioconjugation of protein

Bioconjugation is the process of linkage formation between a biomolecule and another moiety.[1] This moiety can be another biomolecule, a fluorescent dye, a functional handle (i.e. tag) or a synthetic polymer.[2] There are purely biological conjugation methods which utilize enzymes such as sortase, horseradish peroxidase, etc.[3] Various chemical approaches have been discovered for bioconjugations of proteins.[1]

Traditionally, amino acid side chains were utilized as functional handles.[4] Nucleophilic amino acid side chains (Cys, Lys, N-terminus) have been used as targets for the development of several conjugation methods.[4] Thiol is the strongest biological nucleophile, and primary amines are also good nucleophiles. Maleimide and iodoacetamide specifically react with Cys, and succinimide and isothiocyanate selectively react with Lys and *N*-terminus.[4] As *N*-terminus has smaller pK_a than Lys, the reaction mixture pH is also an important factor to the selectivity of the reaction.[5]

After the discovery of an expanded genetic code, non-standard amino acids can be added into the proteins.[6] Incorporated unnatural amino acids can be exploited for bio-orthogonal reactions such as azide-alkyne click reaction,[7] UV crosslinking reaction of benzophenone or aziridine,[8] oxime formation of ketone with hydroxylamine, or hydrazide.[9]

More recently, retro-cycloaddition of tetrazine and *trans*-cyclooctene was added as an option for the unnatural protein modification.[10] For the natural protein modification, Met-selective bioconjugation with oxaziridine reagents was reported based on the redox reactivity of

Met.[11]

1.2 Fluorescence: tagging and biosensor

Fluorescence is an exquisite analytical method for visualizing biomolecules, due to its sensitivity and high signal to noise ratio.[12] Genetically fused green fluorescent protein,[13] and chemically attached fluorophores can visualize a specific biomolecule itself or its activity.[14] Fluorescent small molecules (fluorescein, rhodamine, cyanine, etc) have advantages in their size, since fluorescent proteins are relatively larger (26.0 kDa) than fluorophores. Among several fluorescent molecules, 1,2-dihydropyrrolo[3,4-b]indolizin-3-one has been known as an emission-varying fluorophore. Its photophysical property significantly alters dependent upon C7, and C9 substituents.[15] Exploiting this character, several biosensors (activity of dual-specific phosphatase, lipid-droplet, mitochondria imaging) were made and became the foundation of high-contents screening (autophagy).[15]

In recent years, super-resolution microscopy such as total internal reflection fluorescence microscope (TIRFM), stimulated emission depletion (STED), and stochastic optical reconstruction microscopy (STORM) facilitated the usage of fluorescence for bio-imaging.[16]

1.3 Sulfur(VI) fluoride exchange: SuFEx

Fluorosulfate belongs to the halosulfate family. As fluorine is the worst leaving group among halides, fluorosulfate compounds have better stability than other halosulfates. Fluorosulfate can be synthesized through the reaction of fluorosulfonic acid (HSO_3F) or gaseous sulfuryl fluoride (SO_2F_2).[17] Sulfuryl fluoride is much more stable than sulfuryl chloride, whose substitution chemistry is usually complicated. Therefore,

use of sulfuryl fluoride is a mild way to prepare sulfate compounds.

Sharpless group reported sulfur (VI) fluoride exchange reaction of aryl fluorosulfate ($Ar-OSO_2F$) and aryl silyl ether under catalytic amount of a base such as DBU and BEMP.[18] Aryl fluorosulfate reacts specifically with aryl silyl ether, even in the presence of an aliphatic alcohol, an aliphatic amine or even a thiol.[18] Owing to its chemoselectivity, several polymers were synthesized through SuFEx chemistry.[19]

Recently, imino sulfur oxy difluoride group ($Ar-N=SOF_2$) was unveiled to react with aryl silyl ether more preferentially than aryl fluorosulfate.[20]

1.4 References

[1] N. Stephanopoulos, M. B. Francis, *Nat. Chem. Biol.* **2011**, *7*, 876-884.

[2] J. Kalia, R. T. Raines, *Curr. Org. Chem.* **2010**, *14*, 138-147.

[3] a) W. J. Bradshaw, A. H. Davies, C. J. Chambers, A. K. Roberts, C. C. Shone, K. R. Acharya, *FEBS Journal* **2015**, *282*, 2097-2114. b) F. W. Krainer, A. Glieder, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 1611-1625.

[4] E. Basle, N. Joubert, M. Pucheault, *Chem. Biol.* **2010**, *17*, 213-227.

[5] W. R. Algar, P. E. Dawson, I. L. Medintz, Chemoselective and Bioorthogonal Ligation Reactions: Concepts and Applications, Wiley-VCH Verlag GmbH & Co, **2017**.

[6] A. de Graaf, M. Kooijman, W. E. Hennink, E. Mastrobattista, *Bioconjugate Chem.* **2009**, *20*, 1281-1295.

[7] J. E. Hein, V. V. Fokin, *Chem. Soc. Rev.* **2010**, *39*, 1302-1315.

[8] R. Pelle, N. B. Murphy, *Nucleic. Acids Res.* **1993**, *21*, 2453-2458.

[9] X. Chen, Y. -W. Wu, *Org. Biomol. Chem.* **2016**, *14*, 5417-5439.

[10] K. Kang, J. Park, E. Kim, *Proteom. Sci.* **2017**, *15*, 1-13.

[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] C. A. Combs, H. Shroff, *Curr, Protoc. Neurosci* **2017**, 79, 2.1.1-2.1.25.

[13] S. J. Remington, *Protein Sci.* **2011**, *20*, 1509-1519.

[14] K. M. Dean, A. E. Palmer, *Nat. Chem. Biol.* **2014**, *10*, 512-523.

[15] E. Kim, Y. Lee, S. Lee, S. B. Park, *Acc. Chem. Res.* **2015**, *48*, 538-547.

[16] M. B. Stone, S. A. Shelby, S. L. Veatch, *Chem. Rev.* **2017**, *117*, 7457-7477.

[17] R. Cramer, D. Coffman, *J. Org. Chem.* **1961**, *26*, 4164-4165.

[18] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2014**, *53*, 9430-9448.

[19] J. Dong, K. B. Sharpless, L. Kwisnek, J. S. Oakdale, V. V. Fokin, *Angew. Chem. Int. Ed.* **2014**, *53*, 9466-9470.

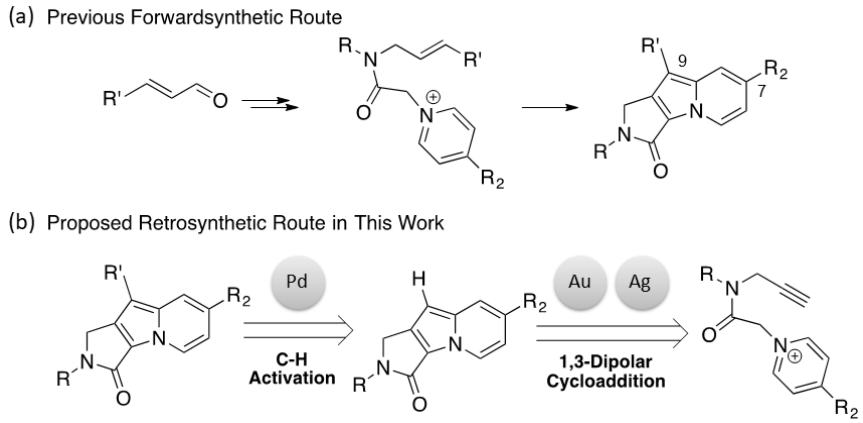
[20] S. Li, P. Wu, J. E. Moses, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2017**, *56*, 2903-2908.

Chapter 2. Developing Synthetic Method of Indolizinone based Fluorescent Molecules

2.1 Introduction

Fluorescence-based sensors for biomolecules have received much attention because of their excellent selectivity and sensitivity, large linear range of analysis, and high spatial resolution through microscopic imaging.[1] Despite the high demand for fluorescent biosensors, the rational design of fluorescent sensors remains challenging because it is difficult to predict changes in the emission wavelength or quantum yield upon a specific recognition event or reaction with analytes. Therefore, an understanding of structure-photophysical-property relationships in fluorophores is needed for the strategic design of novel fluorescent biosensors. Even though there are dozens of reports on correlations between the emission wavelength and chemical structure of fluorophores, such as BODIPY and cyanine dyes,[2] only a few studies have addressed the systematic explanation of fluorescence quantum yields through changes in their substructure.[3]

In relation with this, a systematic approach towards emission-tunable indolizin-based fluorophores has been reported.[4a] Alteration of the substituents at the C7 and C9 positions of the fluorophore resulted in dramatic changes in emission wavelength over a wide range of visible-light wavelengths. Furthermore, the emission wavelength can be predicted by computational analysis of the energy gap between the excited and ground states.[4b] Because of the ability to predict the emission wavelength, several bio-applicable fluorescence sensors and high contents screening system were developed.[5]



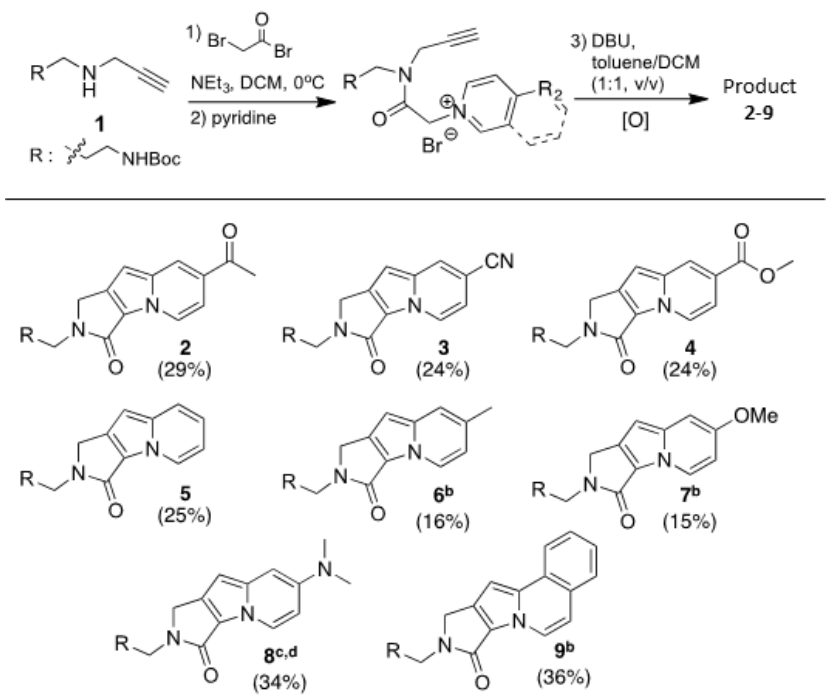
Scheme 2.1. Comparison of a) our previous forward synthetic route and b) proposed retrosynthetic route for the efficient synthesis.

With these studies, we clearly demonstrated the importance of the substituents at the C7 and C9 positions in perturbing the photophysical properties. However, the established synthetic route involves an intramolecular 1,3-dipolar cycloaddition of an azomethine ylide with a cinnamaldehyde-derived olefin. Success of the cycloaddition reaction is significantly influenced by the electronic character of the R^1 and R^2 substituents (Scheme 2.1a). In the case of strongly electron rich R^2 substituents, various attempts have not been successful in affording the desired cycloadducts. The preparation of cinnamaldehydes with electron-rich substituents has also proven difficult. More importantly, the introduction of R^1 substituents at an early stage of the synthetic route was strategically unfavorable for the systematic diversification of fluorescent molecules. For the efficient and facile synthesis, we envisioned a new synthetic route based on a silver or gold-catalyzed intramolecular 1,3-dipolar cycloaddition to prepare the lactam-embedded indolizin-one core and a subsequent palladium-mediated site-specific C-H activation for the late-stage diversification of the fluorophore (Scheme 2.1b).

2.2 Results and Discussion

2.2.1 Synthesis of Indolizin-one through 1,3-Dipolar Cycloaddition

We first introduced terminal alkynes instead of olefins to overcome the limited substrate scope of pyridine-based dipoles and facilitated the 1,3-dipolar cycloaddition of the terminal alkyne with the resulting azomethine ylide by the use of a coinage-metal catalyst.[6] This approach enabled the synthesis of Seoul-Fluors containing diverse R^2 substituents. To optimize the reaction conditions for the lactam embedded indolizin-one core, we prepared the secondary amine **1** through a substitution reaction of propargyl amine with Boc-protected 3-bromopropan-1-amine. The resulting amine **1** was acylated with bromoacetyl bromide and treated with pyridine derivatives to generate pyridinium intermediates (Scheme 2.2).



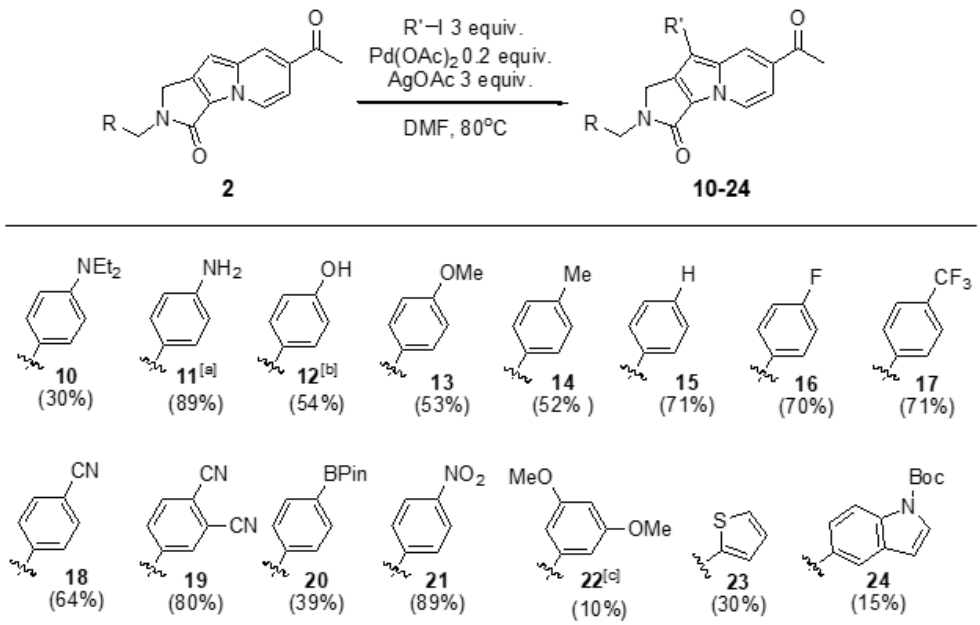
Scheme 2.2. Synthesis of lactam-embedded indolizin-one core structures. The yields given are for the isolated product after four steps. [a] Ag_2O was used to catalyze the 1,3-dipolar cycloaddition. [b] $AuPPh_3Cl$ was used as a catalyst; DDQ was added for the oxidative aromatization.

The tricyclic adducts underwent spontaneous aromatization to give lactam embedded indolizines **2-7** and **9** without the addition of oxidants, such as DDQ, which was required in the previous synthetic method.[4] In the case of the electron-rich 4-(dimethylamino)pyridine, we failed to synthesize the desired indolizine core by the previous method because of the inherently low reactivity of the electron-rich dipole. However, we succeeded in synthesizing the cyclized adduct **8** with a 4-dimethylamino group by a new synthetic route in the presence of chloro(triphenylphosphine)gold (I) as a catalyst. The addition of a catalytic amount of silver (I) oxide also improved the yield of the 1,3-dipolar cycloaddition for pyridines containing weakly electron donating R^2 substituents (**6** and **7**) or isoquinoline (**9**). Therefore, this new synthetic route not only expands the range of possible R^2 substituents, but also improves the efficiency of the synthesis. Under these conditions, we were able to synthesize lactam embedded indolizine cores with various R^2 substituents.

2.2.2 Coupling Reaction of the Fluorophore and Aryl Iodide through Pd-mediated C-H Activation

After the formation of the indolizine core, we aimed to use a transition metal catalyzed site-specific C-H activation[7] at C9 position for the late-stage diversification of the R^1 group. Metal catalyzed C-H

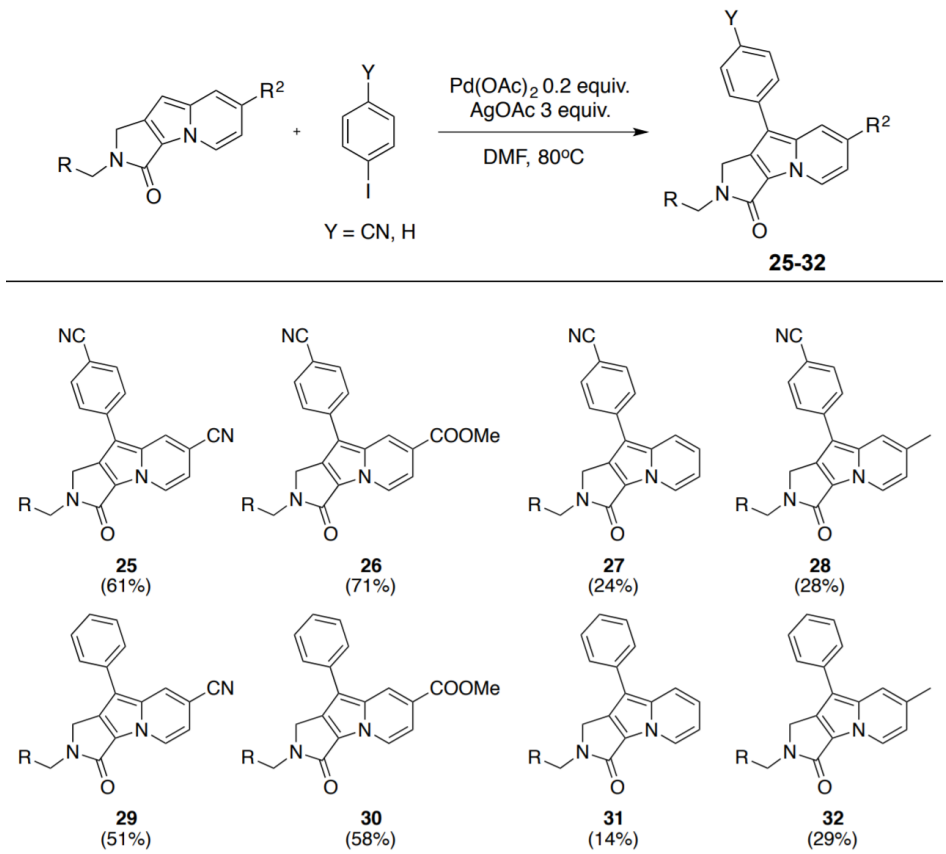
activation methods are powerful, straightforward, and atom economical synthetic tools for carbon-carbon and carbon-heteroatom bond formation.[7a,b] Palladium (Ⅱ) catalyzed C-H activation has emerged as an important catalytic transformation because of its versatility for the installation of many different types of bonds and its exceptional practicality under ambient conditions in the presence of moisture.[7c,d] Therefore, we envisioned the efficient synthesis of those fluorescent molecules containing diverse R^1 groups that were unobtainable by the previously reported route by the cross-coupling of the lactam embedded indolizine cores with aryl iodides.



Scheme 2.3. Late-stage modification of C9 through a cross-coupling reaction involving palladium-mediated C-H activation. The yields given are for the isolated product. [a] Compound **11** was obtained by the reduction of the nitrophenyl compound. [b] Compound **12** was obtained by deprotection of the TBS-protected 4-hydroxyphenyl compound. [c] 3,5-dimethoxyphenyl bromide was utilized as a substrate.

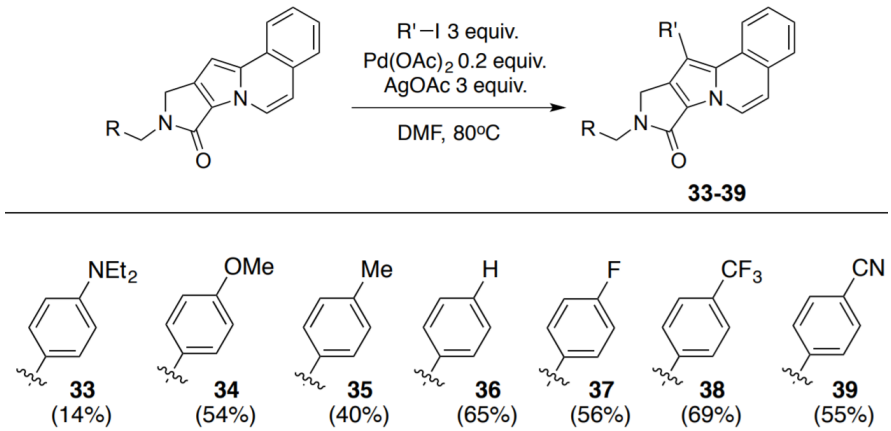
We selected indolizine **2**, which contains an acetyl group at the C7

position, as a model substrate for C-H activation owing to its better stability than other indolizine cores. By screening various catalysts, ligands, oxidants, solvents, and temperatures (results not shown), we found that the optimal cross-coupling conditions for the site-specific C-H activation of indolizine at the C9 position were conventional heating in the presence of palladium (II) acetate and silver acetate.[8] Under these optimized conditions, we successfully synthesized a series of fluorescent molecules with various R^1 substituents, ranging from electron-rich 4-diethylaminophenyl to electron-deficient phthalonitrile, in moderate to good yield, as well as other R^1 substituents, including heterocycles (Scheme 2.3).



Scheme 2.4. Generality of the cross coupling reaction to indolizine cores (**3-6**).

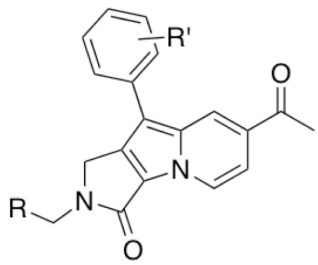
This palladium-mediated cross-coupling reaction also enabled the introduction of R^1 substituents onto other various indolizine cores (compounds **3-6**) containing R^2 substituents with a range of electronic characteristics (Scheme 2.4), and even onto indolizine **9** which has an electron-rich isoquinoline moiety (Scheme 2.5).



Scheme 2.5. Scope of cross coupling reaction of **9**.

2.2.3 Fluorescence Quenching through PeT

Notably, the yield of the palladium catalyzed cross-coupling reaction increased as the electron-donating ability of the R^1 substituent decreased. This tendency was reported in the previous publication about turn-off sensor of phosphatase.[4] We measured the absorption maxima (abs), emission maxima (em), and absolute quantum yields (Φ_F) of compounds **10-19** in methanol. To quantify the electronic character of the R^1 substituents, we used the Hammett substituent constant σ_p, as described previously.[4, 9]



Product	R'	σ_p[a]	E_h[b]	abs[c]	em[d]	Φ_F[e]
10	4-NEt_2	-0.72	-0.160	441	-	0.00
11	4-NH_2	-0.66	-0.169	419	546	0.03
12	4-OH	-0.37	-0.200	413	572	0.12
13	4-OMe	-0.27	-0.195	412	555	0.20
14	4-Me	-0.17	-0.215	404	534	0.38
15	H	0	-0.227	403	523	0.50
16	4-F	0.06	-0.225	402	525	0.47
17	4-CF_3	0.54	-0.249	405	514	0.64
18	4-CN	0.66	-0.247	398	512	0.63
19	3,4-$(CN)_2$	-	-0.262	402	500	0.72

Table 2.1. Electronic effect of R' substituents on the relationship between the structure and photophysical properties in methanol. R represents *N*-Boc-aminoethyl group. [a] σ_p is the Hammett constant of the R'substituent at the *para* position.[9] [b] Calculated energy in hartree of the HOMO level of the phenyl moiety. [c] Longest absorption wavelength maximum. [d] The emission wavelength for excitation at the longest absorption wavelength maximum. [e] Absolute fluorescence quantum yield.

By using density functional theory (DFT), we also calculated the energy level of HOMO in the R^1 group as a barometer of the electron density of the phenyl moiety. The energy gap between the ground and excited states was previously used to predict the emission wavelength, owing to similar patterns of nonradiative energy loss.[4b] However, this energy gap is less well correlated with changes in the quantum yield,

Φ_F. On the other hand, we observed interesting structure–photophysical-property relationships that depended on the electronic character of the R' substituent: Whereas the emission wavelength showed an inverse correlation with the Hammett constant, the quantum yield was positively correlated with the Hammett constant (Table 2.1). In this study, we focused on the quantum-yield variation in the indolizine fluorophore, since a systematic understanding of this property is important for the design of on/off-switchable fluorescent sensors. In fact, Nagano and co-workers reported a mechanistic explanation of fluorescence off/on switching in fluorescein analogues on the basis of phenyl-substituent changes.[3a] Their understanding of quantum yields enabled them to develop fluorescent probes of UDP glucuronosyltransferase 1A1.[10]

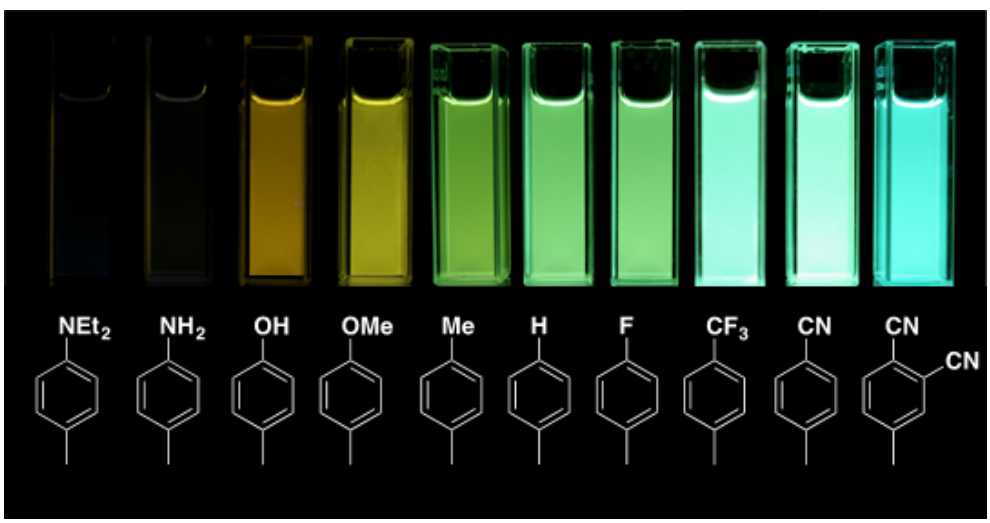


Figure 2.1. Change in the fluorescence brightness of with various R^1 substituents.

We observed a systematic decrease in the quantum yield from 0.72 (for **19**) to 0.00 (for **10**) as the electron density of the phenyl group was altered by changing the R^1 substituent from a dicyano-substituted to a 4-diethylamino-substituted phenyl moiety (Figure 2.1). Similar

fluorescence quenching behavior, correlated with the electron richness of the R^1 substituent, was also observed in three different solvents: methanol, acetonitrile, and dichloromethane. This dramatic change in the quantum yield can be rationalized by the occurrence of photoinduced electron transfer (PeT), one of the major mechanisms of fluorescence quenching, through electron transfer from a PeT donor to its acceptor in the fluorescent molecule.[3, 11] Therefore, we considered the phenyl part as a PeT donor and their indolizine core as a PeT acceptor. The quantum yields were correlated with two major parameters to enable quantification of the electron-donating characteristics of the R' substituents: the Hammett substituent constant (Figure 2.2a) and the HOMO energy level (Figure 2.2b) of the phenyl moiety.

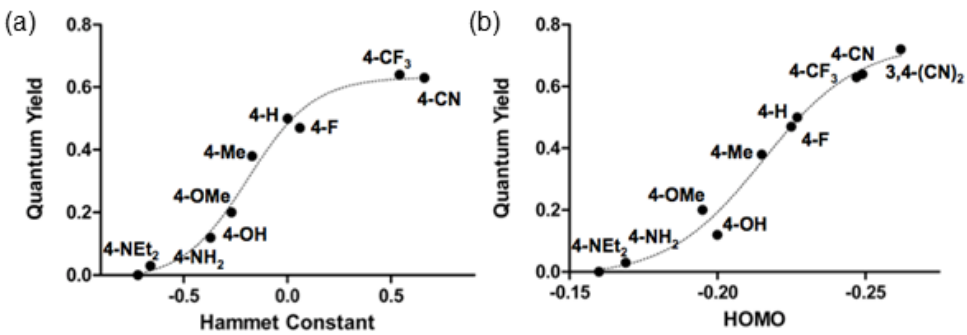


Figure 2.2 Relationship between the quantum yield with (a) Hammett substituent constant and (b) HOMO energy level of the phenyl moiety.

To confirm the PeT process, we tried to detect the existence of radical species transiently generated upon electron transfer from the PeT donor to the acceptor. By using EPR spectroscopy, we directly observed radical signals in the nonfluorescent compound **10** under irradiation at 365 nm as a signature event of the PeT process (Figure 2.3 upper). However, no EPR signal was observed from the fluorescent **15**.

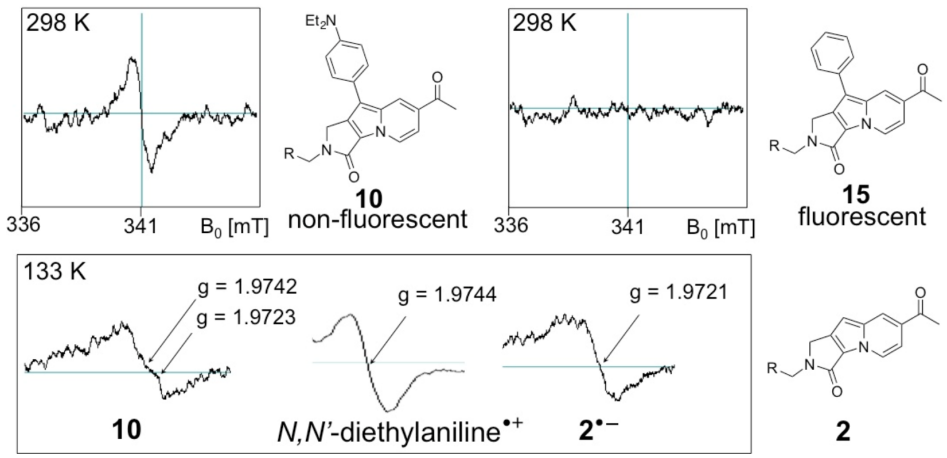
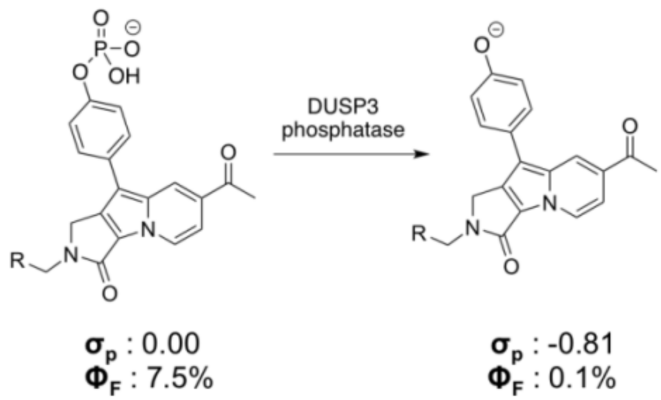


Figure 2.3. Chemical structures of **2**, **10**, and **15** and their EPR spectra (1.0 mM) measured in acetonitrile at room temperature and 133 K under irradiation at 365 nm with a high-intensity xenon lamp.

In fact, the spectrum of compound **10** observed at 133 K is composed of two signals due to a radical cation (g = 1.9742) and a radical anion (g = 1.9723), whereas the g values of [N,N-diethylaniline]$^{\cdot+}$ and **2**$^{\cdot-}$ were 1.9744 and 1.9721, respectively (Figure 2.3). This observation confirmed that the electron-richness of the R^1 moiety can induce the PeT process, which affects the quantum yield. On the basis of our understanding of the structure-quantum-yield relationship, we were able to rationally explain the fluorescence of on/off-switchable sensors. For example, previously reported fluorescence turn-off bioprobe, S*f*BP, as a dual-specific phosphatase DUSP3. Upon the selective dephosphorylation of S*f*BP by DUSP3, the quantum yield drastically decreased from 7.5 to 0.1%. This 75-fold quantum-yield r eduction in S*f*BP can be rationalized by differences in the electron richness of the phosphate group (σ_p = 0.0) and the phenoxide group (σ_p = 0.81) upon a specific chemical transformation of S*f*BP by DUSP3 (Scheme 2.6).



Scheme 2.6. Correlation between quantum yields and Hammet constants of dual specific phosphatase bioprobe, S*f*BP, and its dephosphorylated counterpart (**12**).

2.2.4 Synthesis of Hydrogen Peroxide Sensor

The rational predictability of quantum-yield changes enables the design and development of fluorescent on/off-switchable sensors. Therefore, we substantiated a rational designing ability by the development of a fluorescent H_2O_2 sensor **20** containing a phenyl boronic acid pinacol ester substituent.[12] The boronate ester in biosensor **20** for reactive oxygen species was transformed into a phenol group upon treatment with H_2O_2 in a aqueous system to give **12** (Figure 2.4).

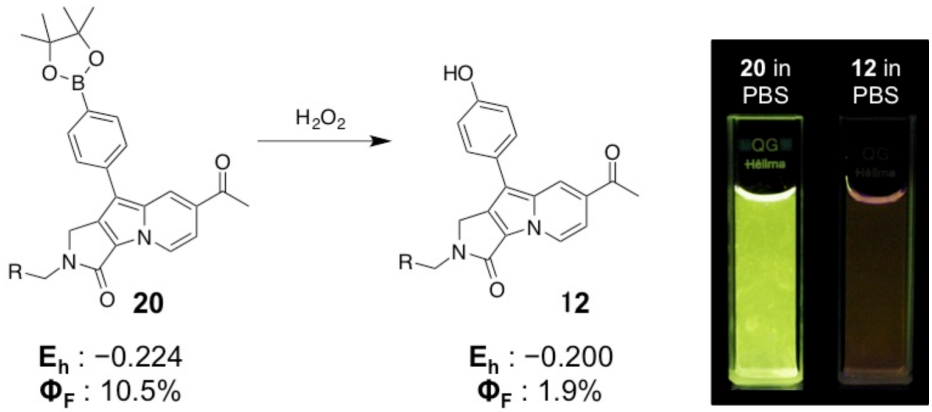


Figure 2.4. Structural change from **20** to **12** and change in the quantum yield upon the addition of H_2O_2. Photographic images of **20** and **12** in phosphate-buffered saline.

The drastic decrease in the quantum yield from 10.5 to 1.9% was able to be predicted on the basis of the change in both the HOMO energy (from 0.224 to 0.200 hartree) and the estimated Hammett constant (from 0.3 to 0.12). We also demonstrated that this turn-off sensor **20** can report real-time changes in reactive oxygene species levels of HeLa human cervical cancer cells (Figure 2.5). The development of a cellular H_2O_2 sensor implies that our study of the structure-quantum-yield relationship has provided a rational design strategy for the construction of novel fluorescence on/off-switchable sensors.

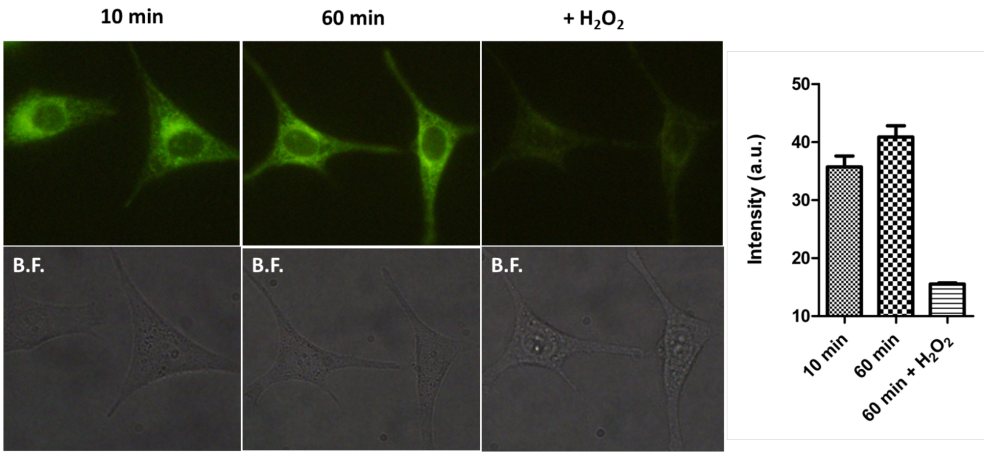
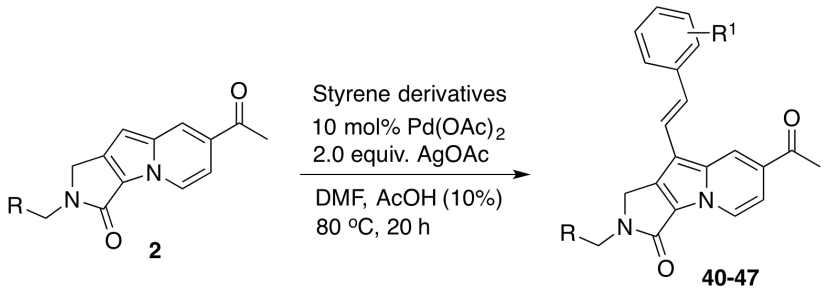


Figure 2.5. Fluorescence intensity of **20** after cellular uptake for 10, 60 min, 1% hydrogen peroxide addition after 60 min at 5% CO_2 incubator at 37 ^{o}C.

2.2.5 Heck type Coupling Reaction of the Fluorophore and Styrene Derivatives

For further extension of the π-conjugation with an additional olefin unit, olefin insertion between phenyl group and fluorophore was done through heck-type C-H activation. Compound **2** was chosen and underwent the reaction. Various substituted styrenes containing both electron-withdrawing (NO_2, F_5, CF_3, Br) and electron-donating functional groups (Me, OMe, NH_2). All of the compounds were obtained in moderate to excellent yields. No regioisomers were observed for any of the styrene derivatives, further confirming the excellent regioselectivity of this direct C-H activation.



Product	R^1	yield	abs[b]	em[c]	Φ_F[d]
40	4-NO_2	83[a]	450	618	0.40
41	2,3,4,5,6-F_5	99	415	527	0.42
42	4-CF_3	76	428	533	0.35
43	4-Br	98	433	545	0.29
44	H	94	424	539	0.30
45	4-Me	42	428	545	0.21
46	4-OMe	49	433	557	0.10
47[e]	4-NH_2	51	450	588	0.10

Table 2.2. Electronic effects of substituents on the isolated yield and photophysical properties. R represents *N*-Boc-aminoethyl group. [a] Recrystalization yield. [b] Longest absorption wavelength maxima. [c] Excited at the longest absorption wavelength maxima. [d] Absolute fluorescence quantum yield. [e] Obtained via the reduction of **40**.

The quantum yield is elaborately controlled by the electronic characters of the C9 substituents, consistent with the previous results (chapter 2.2.3). Increasing the electron-donating ability of the R^1 substituent in styryl analogues from F_5 (**41**) to NH_2 (**47**) produced a bathochromic shift in the emission wavelength from 527 nm to 588 nm. The emission wavelength of a compound with NO_2 (**40**) shows a significantly red-shifted emission wavelength (618 nm) compared to others. This phenomenon was presumably due to the characteristics of the nitro group via the $n \rightarrow \pi^*$ transition.

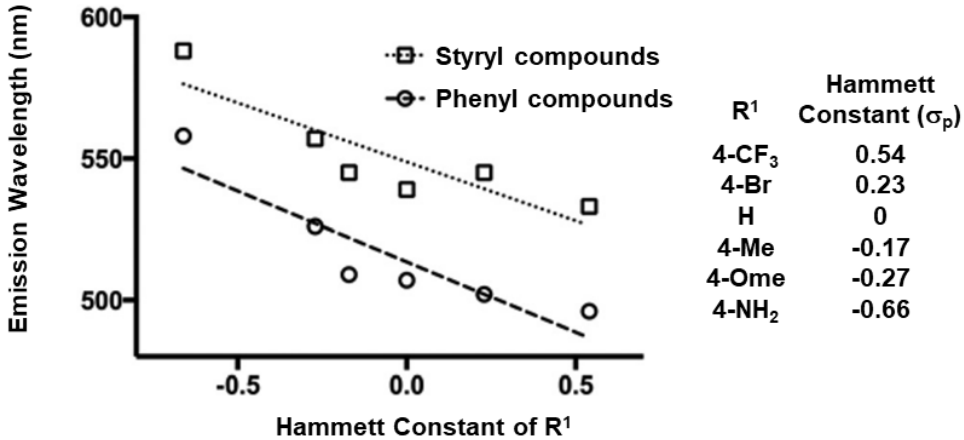


Figure 2.6. Correlation plot between emission wavelength and Hammet constant of R^1 substituents.

The emission wavelength of styryl analogues showed an average of 39 nm of bathochromic shift compared to that of the original ones (Figure 2.6.). This result confirms that the extension of the π-conjugated system by olefin insertion between the indolizine core and C9 substituents, leads to the perturbation of its photophysical properties.

2.3 Conclusion

In this chapter, we developed an efficient method for the facile synthesis of indolizin-one based fluorescent molecules by the silver- and gold catalyzed intramolecular 1,3-dipolar cyclization of terminal acetylenes with substituted azomethine ylides and subsequent palladium-mediated C-H activation for late-stage diversification. By using this new synthetic route, we were able to synthesize diverse fluorescent molecules containing a wide variety of R^1 and R^2 substituents, that were unobtainable including styryl groups. And we discovered that their quantum yield can be perturbed by systematic changes in the electronic character of the R' substituents, as quantified by Hammett constants and the HOMO energy level of the R^1 group. Furthermore, we demonstrated the occurrence of a photoinduced electron-transfer process might cause a reduction in the quantum yield of the indolizinone fluorophore with increasing electron richness of the R^1 substituent. By using our fundamental understanding of the structure -photophysical property relationship, we designed and developed a turn-off chemosensor for hydrogen peroxide; this compound was not accessible by our previous synthetic route. Therefore, we envision the development of novel fluorescent on/off-switchable sensors by rational design with respect to PeT-based changes in quantum yield. Furthermore, with the new synthetic route in hand, we could efficiently access a diverse collection of styryl analogues with excellent substrate generality via the late-stage diversification. We observed the general bathochromic shifts of emission wavelength maxima in styryl analogues (39 nm on average).

2.4 Experimental Section

2.4.1 Materials and Methods

^{1}H and ^{13}C NMR spectra were recorded on Bruker DRX-300 (Bruker Biospin, Germany), Agilent 400-MR (Agilent Technologies) and Varian Inova-500 (VarianAssoc, Palo Alto, USA), chemical shifts were measured in ppm downfield from internal tetramethylsilane (TMS) standard. Multiplicity was indicated as follows: s (singlet); d (doublet); t (triplet); q (quartet); m (multiplet); dd (doublet of doublet); dt (doublet of triplet); br s (broad singlet), etc. Coupling constants were reported in Hz. Low resolution mass analyses were performed on LC/MS system, Finnigan MSQ plusSurveyer (Thermo) equipped with a reverse phase column (C-18, 50 × 2.1 mm, 5 μm) and photodiode array detector using electron spray ionization (ESI) or performed on 6120 Quadrupole LC/MS (Agilent Technologies). The identity of desired compounds were further confirmed by high-resolution mass spectrometry (HRMS). The HRMS analyses were conducted at the research analysis center of Korea Advanced Institute of Science and Technology by direct injection on a Bruker Daltonik micro TOF-Q2 spectrometer using electron spray ionization (ESI). Absorbance of final fluorescent compounds were measured by UV-VIS spectrophotometer UV-1650 PC (Shimatzu,Japan). Excitation and Emission maxima were measured by Cary Eclipse Fluorescence spectrophotometer (Varian Assoc., Palo Alto, USA). Absolute quantum yield was measured by absolute PL quantum yield measurement system QE-1000 (OTSUKA Electronics). The ESR measurements of the photo-excited compound were carried out with electron spin resonance (ESR) spectrometer JES-TE200 (JEOL) equipped with a variable-temperature apparatus to detect the transient radical species in a solution of sample (1 mM) in acetonitrile at 298 K and 133

K under 365-nm irradiation with a high intensity xenon lamp B-100AP (UVP). p-phenylene diamine and $[Ru(bpy)_3](PF_6)_3$ were used to make an authentic radical of 2 and *N,N*-diethylaniline. All the photography images of the fluorescent compounds (final concentration = 20 μM) were taken under irradiation with 365nm wavelength. *In silico* calculations were performed using the Materials Studio® 4.2 program (AccelrysSoftware Inc.) A generalized gradient approximation (GAA) for the exchange correlation function of Perdew, Burke, and Ernzerhof (PBE) was used with the double numerical basis set with polarization (DNP) as implemented in DMol3.

All chemical reagents including 3-bromopropylamine hydrobromide, di-*tert*-butyldicarbonate, propargyl amine, triethylamine, DBU, bromoacetyl bromide, pyridine derivatives, DDQ, iodobenzene derivatives, palladium acetate, silver oxide, silver acetate, chloro(triphenylphosphine)gold, anhydrous dimethyl formamide, *N,N*-diethylaniline and *p*-phenylene diamine were purchased either from Sigma-Aldrich, Tokyo Chemical Industry Co., or Acros, and used without further purification. $[Ru(bpy)_3](PF_6)_3$ was obtained by previously reported method.[13] The progress of reaction was monitored using thin-layer chromatography (silica gel 60 F_{254} 0.25 mm), and components were visualized by observation under UV light (254 and 365 nm) or by treating TLC plates with anisaldehyde or ninhydrin, followed by heating. Solvents were purchased from commercial venders and used without further purification. Distilled water was polished by ion exchange and filtration. Sodium acetate buffer solution pH 4.5 was prepared by adjusting the pH after solvation of 63 g of anhydrous sodium acetate in doubly deionized water and 90 mL of acetic acid, and dilute to 1000 mL of water.

We carried out fluorescence bioimaging studies with Olympus Inverted

MicroscopeModel IX71, equipped for epi-illumination using a halogen bulb (Philips No. 7724). Emission signal of each experiments were observed using a 450-480 band pass exciter filter, a 500 nm center wavelength chromatic beam splitter, a 515 nm-long pass barrier filter (Olympus filter set U-MWB2). Emission signal of each experiments were detected with 12.5M pixel recording digital color camera (Olympus, DP71). HeLa cells were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). HeLa human cervical cancer cell lines were cultured in RPMI 1640 (GIBCO, Invitrogen) supplemented with heat-inactivated 10% (v/v) fetal bovineserum (GIBCO, Invitrogen) and 1% (v/v) antibiotic-antimycotic solution (GIBCO, Invitrogen). Cells were maintained in a humidified atmosphere of 5 % CO2 incubator at 37 °C, and cultured in 100-mm cell culture dish (CORNING).

2.4.2 Preparations of Compounds

Preparation of tert-butyl (3-(prop-2-yn-1-ylamino)propyl)carbamate (1): 3-Bromopropylamine hydrobromide (5 g, 22.5 mmol) was added to a solution of Di-*t*-butyl dicarbonate (5 g, 22.5 mmol) and DIPEA (10 mL) in DCM (80 mL). After the reaction completion monitored by TLC, the solvent was evaporated and diluted in acetonitrile (80 mL). Solution was heated to 60 °C and propargyl amine (5.8 mL, 90 mmol) was added. After the reaction completion monitored by TLC, solvent was evaporated and diluted in DCM (30 mL). Saturated $NaHCO_3$ was added to the solution and the organic material was extracted 2 times with DCM. Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4, and concentrated in vacuo after filteration. The

residue was purified by silica gel flash column chromatography to afford 4.02 g of clear oil (84%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.97 (br s, 1H), 3.42 (d, J = 3.0Hz, 2H), 3.20 (dd, J = 12.0, 4.0 Hz, 2H), 2.74 (t, J = 7.5 Hz, 2H), 2.21 (t, J = 3.0 Hz, 1H), 1.70-1.61 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 156.1, 82.0, 71.4, 46.2, 38.1, 29.8, 28.4; LRMS (ESI+) *m/z* calcd for $C_{11}H_{20}N_2O_2$ $[M+H]^+$ 213.16; found : 213.11.

Preperation of 2-(4-iodophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane: 2-(4-iodophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane was synthesized by previously reported method.[14]

General synthetic procedure for 1,2-dihydro-3H-pyrrolo [3,4-b]indolizin-3- one derivatives: Bromoacetyl bromide (1.95 mL, 22.7 mmol) was dissolved in dichloromethane (80 mL) was cooled down to 0 °C for 10 min. To this solution, a mixture of tert-butyl(3-(prop-2-yn-1-ylamino)propyl)carbamate (**1**) (4.02 g, 18.9 mmol) and triethylamine (8 mL, 57 mmol) in dichloromethane (20 mL) was slowly added. Then, the reaction mixture was gradually warmed up to room temperature and stirred for 1 h. When the reaction was completed, saturated NH_4Cl (aq) was added to the reaction mixture and the organic material was extracted twice with dichloromethane. Combined organic layer was washed with brine, dried over anhydrous Na_2SO_4 (s), and concentrated in vacuo after filtration. The residue was purified by silica-gel flash column chromatography to afford 6.7 g of tert-butyl(3-(2-bromo-N-(prop-2-yn-1-yl)acetamido)propyl)carbamate as a clear oil (89%). After addition of pyridine derivatives (1 equiv.) to a solution of the resulting intermediate in dichloromethane (1 M), the reaction mixture was heated at 60 °C. After the reaction completion, the reaction mixture was diluted with dichloromethane/toluene (1:1, v/v, 0.02 M), added with DBU

(2 equiv.) and stirred for overnight at 80 °C. In case of compounds **6**, **7**, and **9**, silver oxide was added during the intramolecular [3+2] cycloaddition. For compound **8**, chloro(triphenylphosphine)gold was added during the cyclization, followed by treatment of DDQ for oxidative aromatization. When the reaction was completed, the reaction mixture was concentrated in vacuo after short filtration with silica S5 gel. The residue was purified by silica-gel flash column chromatography to afford the desired product (**2-9**).

tert-Butyl (3-(7-acetyl-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (2): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.52 (d, J = 9.0 Hz, 1H), 8.12 (s, 1H), 7.25 (d, J = 7.5 Hz, 1H), 6.65 (s, 1H), 5.39 (br s, 1H), 4.36 (s, 2H), 3.66 (t, J = 7.5 Hz, 2H), 3.16 (dd, J = 13.5, 7.5 Hz, 2H), 2.62 (s, 3H), 1.87-1.78 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.6, 161.9, 156.1, 138.0, 136.6, 128.7, 124.4, 122.1, 109.1, 98.5, 79.1, 46.5, 40.0, 37.3, 28.8, 28.4, 26.0; HRMS (ESI+) *m/z* calcd for $C_{20}H_{25}N_3O_4$ $[M+Na]^+$ 394.17; found: 394.1724.

tert-Butyl (3-(7-cyano-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (3): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.58 (d, J = 6.0 Hz, 1H), 7.90 (s, 1H), 6.81 (d, J = 9.0 Hz, 1H), 6.65 (s, 1H), 5.31 (br s, 1H), 4.39 (s, 2H), 3.67 (t, J = 6.0 Hz, 2H), 3.18 (dd, J = 13.5, 7.5 Hz, 2H), 1.89-1.80 (m, 2H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 161.6, 156.1, 137.1, 137.0, 126.2, 125.2, 123.7, 118.3, 111.0, 102.7, 98.0, 79.1, 46.5, 40.1, 37.3, 28.8, 28.4; HRMS (ESI+) *m/z* calcd for $C_{19}H_{22}N_4O_3$ $[M+Na]^+$ 377.16; found: 377.1570.

Methyl 2-(3-((tert-butoxycarbonyl)amino)propyl)-3-oxo-2,3-dihydro

-1Hpyrrolo[3,4-b]indolizine-7-carboxylate (4): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.52 (d, J = 6.0 Hz, 1H), 8.25 (s, 1H), 7.25 (d, J = 6.0 Hz, 1H), 6.60 (s, 1H), 5.40 (br s, 1H), 4.35 (s, 2H), 3.94 (s, 3H), 3.65 (t, J = 6.0 Hz, 2H), 3.16 (dd, J = 12.0, 6.0 Hz, 2H), 1.86- 1.78 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 165.9, 162.0, 156.1, 138.3, 136.6, 124.2, 122.8, 121.5, 110.3, 100.0, 97.7, 79.0, 52.3, 46.4, 40.0, 37.3, 28.8, 28.4; HRMS (ESI+) *m/z* calcd for $C_{20}H_{25}N_3O_5$ $[M+Na]^+$ 410.17; found: 410.1679.

tert-Butyl (3-(3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2- yl)propyl) carbamate (5): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.52 (d, J = 6.0 Hz, 1H), 7.43 (d, J = 9.0 Hz, 1H), 6.90 (dd, J = 18.0, 9.0 Hz, 1H), 6.67 (dd, J= 15.0, 9.0 Hz, 1H), 6.32 (s, 1H), 5.54 (br s, 1H), 4.29 (s, 2H), 3.62 (t, J = 7.5 Hz, 2H), 3.15 (dd, J = 13.5, 7.5 Hz, 2H), 1.84-1.75 (m, 2H), 1.42 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.6, 156.2, 139.9, 136.4, 125.0, 120.4, 120.2, 119.2, 111.3, 93.3, 78.9, 46.4, 39.8, 37.2, 28.8, 28.4; HRMS (ESI+) *m/z* calcd for $C_{18}H_{23}N_3O_3$ $[M+Na]^+$ 352.16; found: 352.1612.

tert-Butyl(3-(7-methyl-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate(6): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.42 (d, *J* = 9.0 Hz,1H), 7.18 (s, 1H), 6.51 (d, *J* = 6.0 Hz, 1H), 6.19 (s, 1H), 5.53 (br s,1H), 4.27 (s, 2H), 3.61 (t, *J* = 7.5 Hz, 2H), 3.15 (dd, *J* = 12.0, 6.0 Hz, 2H), 2.33 (s, 3H), 1.84-1.74 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.6, 156.2, 140.3, 136.8, 130.9, 124.4,117.4, 114.0, 109.0, 91.9, 78.9, 46.4, 39.7, 37.2, 28.9, 28.4, 21.3; HRMS (ESI+) *m/z* calcd. for $C_{19}H_{25}N_3O_3$ $[M+Na]^+$ 366.18; found: 366.1766.

tert-Butyl(3-(7-methoxy-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate(7): ^{1}H NMR (300MHz, $CDCl_3$) δ 8.34 (d, *J* = 9.0 Hz, 1H), 6.66 (d, *J* = 3.0 Hz, 1H), 6.38 (dd, *J* = 9.0, 3.0 Hz, 1H), 6.12 (s,

1H), 5.55 (br s, 1H), 4.23 (s, 2H), 3.80 (s, 3H), 3.58 (t, 6.0 Hz, 2H), 3.13 (dd, *J* = 12.0, 6.0 Hz, 2H), 1.80-1.71 (m, 2H), 1.41 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.5, 154.2, 144.1, 137.8, 126.0, 118.8, 106.4, 100.0, 95.9, 91.6, 55.2, 46.4, 39.7, 37.2, 28.9, 28.4; HRMS (ESI+) *m/z* calcd. for $C_{19}H_{25}N_3O_4$ $[M+Na]^+$ 382.17; found: 382.1719.

tert-Butyl (3-(7-(dimethylamino)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b] indolizin -2-yl)propyl)carbamate (8): ^{1}H NMR (300 MHz, DMSO) δ 8.14 (d, *J*= 8.0 Hz, 2H), 7.06-7.02 (m, 2H), 5.26 (s, 2H), 4.28 (br s, 2H), 4.16 (br s, 2H), 3.21 (s, 6H), 1.84-1.79 (m, 2H), 1.38-1.36 (m, 9H); LRMS (ESI-) *m/z* calcd. for $C_{20}H_{28}N_4O_3$ $[M-H]^-$ 371.21; found: 371.79.

tert-Butyl(3-(8-oxo-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo[2,1-a]isoquino lin-9-yl)propyl)carbamate(9): ^{1}H NMR (300MHz, $CDCl_3$) δ 8.27 (d, *J* = 9.0 Hz, 1H), 7.96 (d, *J* = 9.0 Hz, 1H), 7.56 (d, *J* = 9.0 Hz, 1H), 7.48-7.36 (m, 2H), 6.82 (d, *J* = 6.0 Hz, 1H), 6.75 (s, 1H), 5.56 (br s, 1H), 4.23 (s, 2H), 3.60 (t,*J* = 6.0Hz, 2H), 3.15 (d, *J* = 6.0 Hz, 2H), 1.83-1.73 (m, 2H), 1.43(s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.5, 156.2, 136.8, 135.2, 127.6, 127.4, 127.2, 126.9, 122.8, 122.7, 122.6, 111.6, 93.9, 78.9, 46.5, 42.7, 39.8, 37.3, 28.9, 28.5; HRMS (ESI+) *m/z* calcd. for $C_{22}H_{25}N_3O_3$$[M+Na]^+$ 402.18; found: 402.1777.

General synthetic procedure of the cross-coupling reaction: To a solution of γ-lactam embedded indolizines (**2**, **3**, **4**, **5**, **6**, **9**) in DMF, iodobenzene derivatives (3 equiv.), palladium acetate (0.2 equiv.) and silver acetate (3 equiv.) were added and stirred at 80 °C for overnight. After the reaction completion as monitored by TLC, reaction mixture was concentrated *in vacuo* after filtration. The residue was purified by silica gel flash column chromatography to afford the desired product

(**10-20**). In case of *tert*-butyl (3-(7-acetyl-9-(4-aminophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate(**11**), 4-nitro iodobenzene was used as a substrate of cross-coupling reaction, and reduced by hydrogenation under H_2, Pd/C condition. In case of *tert*-butyl(3-(7-acetyl-9-(4-hydroxyphenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate(**12**), TBS-protected 4-hydroxy iodobenzene was used as a substrate of cross-coupling reaction, and deprotectedin tetrabutylammonium fluoride solution.

***tert*-Butyl (3-(7-acetyl-9-(4-(diethylamino)phenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (10)**: ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.47 (d, J = 8.0 Hz, 1H), 8.36 (s, 1H), 7.38 (d, J = 8.0 Hz, 2H), 7.21 (d, J = 8.0 Hz, 1H), 6.79 (d, J = 8.0 Hz, 2H), 5.39 (br s, 1H), 4.45 (s, 2H), 3.66 (t, J = 6.0 Hz, 2H), 3.41 (q, J = 8.0 Hz, 4H), 3.15 (dd, J = 12.0, 4.0 Hz, 2H), 2.58 (s, 3H), 1.84-1.76 (m, 2H), 1.41 (s, 9H), 1.21 (t, J = 8.0 Hz, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 195.5, 162.0, 146.7, 134.02, 133.95, 128.6, 128.3, 124.4, 122.0, 120.0, 115.0, 112.2, 109.1, 91.9, 77.2, 46.5, 44.4, 40.0, 37.2, 28.7, 28.4, 26.0, 12.6; HRMS (ESI+) *m/z* calcd. for $C_{30}H_{38}N_4O_4$ $[M+Na]^+$ 541.28; found: 541.2812.

***tert*-Butyl (3-(7-acetyl-9-(4-aminophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (11)**: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.54 (d, J = 6.0 Hz, 1H), 8.37 (s, 1H), 7.37 (d, J = 9.0 Hz, 2H), 7.26 (d, J = 6.0 Hz, 1H), 6.85 (d, J = 9.0 Hz, 2H), 5.38 (br s, 1H), 4.,49 (s, 2H), 3.70 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 2.62 (s, 3H), 1.90-1.80 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.6, 161.9, 156.1, 145.7, 134.20, 134.17, 128.6, 128.5, 124.4, 123.4, 122.4, 121.6, 115.7, 114.7, 109.2, 79.1, 46.5, 40.1, 37.3, 29.7, 28.4, 26.0; HRMS (ESI+) *m/z* calcd. for $C_{26}H_{30}N_4O_4$ $[M+Na]^+$ 485.22; found: 485.2170.

tert-Butyl(3-(7-acetyl-9-(4-hydroxyphenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (12): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.48 (d, J = 6.0 Hz, 1H), δ 8.29 (s, 1H), 7.33 (d, J = 6.0 Hz, 2H), 7.21 (d, J = 6.0 Hz, 1H), 6.97 (d, J = 9.0 Hz, 2H), 5.56 (br s, 1H), 4.42 (s, 2H), 3.65 (t, J = 6.0 Hz, 2H), 3.17 (d, J = 6.0 Hz, 2H), 2.57 (s, 3H), 1.87 -1.78 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.9, 162.1, 156.1, 134.5, 134.4, 128.74, 128.66, 124.7, 124.4, 122.3, 121.5, 116.3, 114.5, 109.4, 79.4, 49.9, 49.6, 46.6, 40.3, 29.7, 28.4, 26.0; HRMS (ESI+) *m/z* calcd. for $C_{26}H_{29}N_3O_5$ $[M+Na]^+$ 486.20; found: 486.2012.

tert-Butyl(3-(7-acetyl-9-(4-methoxyphenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (13): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.54 (d, J = 9.0 Hz, 1H), 8.36 (s, 1H), 7.49 (d, J = 9.0 Hz, 2H), 7.27 (d, J = 9.0 Hz, 1H), 7.07 (d, J = 9.0 Hz, 2H), 5.38 (br s, 1H), 4.49 (s, 2H), 3.90 (s, 3H), 3.70 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 2.62 (s, 3H), 1.90-1.80 (m, 2H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.5, 158.7, 156.1, 134.5, 134.4, 128.9. 128.8, 125.9, 124.5, 122.6, 121.3, 114.8, 109.4, 79.1, 55.4, 46.5, 40.1, 29.7, 28.4, 26.1; HRMS (ESI+) *m/z* calcd. for $C_{27}H_{31}N_3O_5[M+K]^+$ 516.19; found: 516.2125.

tert-Butyl(3-(7-acetyl-3-oxo-9-(p-tolyl)-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (14): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.53 (d, J = 9.0 Hz, 1H), 8.38 (s, 1H), 7.45 (d, J = 6.0 Hz, 2H), 7.33 (d, J = 6.0 Hz, 2H), 7.26 (d, J = 6.0 Hz, 1H), 5.39 (br s, 1H), 4.49 (s, 2H), 3.69 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 2.61 (s, 3H), 2.45 (s, 3H), 1.89-1.81 (m, 2H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.5, 161.9, 156.1, 136.8, 134.7, 134.5, 130.5, 130.0, 129.0, 128.2, 127.4, 124.5, 122.6, 121.4, 114.2, 109.5, 79.1, 46.5, 40.1, 37.4, 28.8, 28.4, 26.0, 21.2;

HRMS (ESI+) m/z calcd. for $C_{27}H_{31}N_3O_4$ $[M+Na]^+$484.22; found: 484.2220.

tert-Butyl(3-(7-acetyl-3-oxo-9-phenyl-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (15): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.55 (d, J = 6.0 Hz, 1H), 8.40 (s, 1H), 7.57-7.49 (m, 4H), 7.37 (t, J = 6.0 Hz, 1H), 7.28 (d, J = 6.0 Hz, 1H), 5.38 (br s, 1H), 4.52 (s, 2H), 3.70 (t, J = 6.0 Hz, 2H), 3.19 (dd, J =12.0, 6.0 Hz, 2H), 2.62 (s, 3H), 1.90-1.82 (m, 2H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.5, 161.8, 156.1, 134.9, 134.6, 133.5, 129.34, 129.25, 127.5, 127.0, 124.6, 122.7, 121.2, 114.1, 109.6, 79.1, 46.6, 40.2, 37.4, 29.7, 28.4, 26.1; HRMS (ESI+) *m/z* calcd. for $C_{26}H_{29}N_3O_4$ $[M+K]^+$ 486.18; found: 486.1998.

tert-Butyl (3-(7-acetyl-9-(4-fluorophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (16): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.54 (d, J = 6.0 Hz, 1H), 8.32 (s, 1H), 7.51 (br t, J = 6.0 Hz, 2H), 7.28-7.18 (m, 3H), 5.35 (br s, 1H), 4.49 (s, 2H),3.69 (t, J = 6.0 Hz, 2H), 3.19 (d, J = 6.0 Hz, 2H), 2.61 (s, 3H), 1.90-1.81 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.4, 161.79 (d, $^1J_{C,F}$ = 246 Hz), 161.75, 156.1, 134.8, 134.6, 129.6, 129.3, 129.2, 129.1, 124.6, 122.7, 120.8, 116.36 (d, $^2J_{C,F}$ = 21.8 Hz), 113.1, 109.7, 79.1, 46.4, 40.2, 37.3, 29.7, 28.4, 26.1; HRMS (ESI+) *m/z* calcd. for $C_{26}H_{28}FN_3O_4$ $[M+K]^+$ 504.17; found: 504.1923.

tert-Butyl (3-(7-acetyl-3-oxo-9-(4-(trifluoromethyl)phenyl)-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (17): ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.57 (d, J = 6.0 Hz, 1H), 8.37 (s, 1H), 7.76 (d, J = 9.0 Hz, 2H), 7.66 (d, J = 9.0 Hz, 2H), 7.30 (d, J = 6.0 Hz, 1H), 5.36 (br s, 1H), 4.53 (s, 2H), 3.70 (t, J = 4.5 Hz, 2H), 3.19 (d, J = 6.0 Hz, 2H), 2.62 (s, 3H), 1.89-1.81 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 195.3, 161.6, 156.1, 137.2, 135.2, 134.9, 129.9, 128.67 (q, $^2J_{C,F}$ = 32.5 Hz,

1H), 127.6, 127.5, 126.25 (q, $^{3}J_{C,F}$ = 3.3 Hz, 1H), 124.8, 124.1 (q, $^{1}J_{C,F}$ = 271 Hz, 1H), 123.0, 120.5, 112.4, 110.0, 79.1, 46.5, 40.2, 37.3, 29.7, 28.7, 28.4, 26.2; HRMS (ESI+) *m/z* calcd. for $C_{27}H_{28}F_3N_3O_4$[M+Na]$^+$ 538.19; found: 538.1934.

***tert*-Butyl (3-(7-acetyl-9-(4-cyanophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo [3,4-b]indolizin-2-yl)propyl)carbamate (18)**: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.61 (d, *J* = 6.0 Hz, 1H), 8.40 (s, 1H), 7.79 (d, *J* = 9.0 Hz, 2H), 7.66 (d, *J* = 6.0 Hz, 2H), 7.34 (d, *J* = 6.0 Hz, 1H), 5.28 (br s, 1H), 4.55 (s, 2H), 3.71 (t, *J* = 7.5 Hz, 2H), 3.19 (dd, *J* = 12.0, 6.0 Hz, 2H), 2.65 (s, 3H), 1.91-1.80 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.3, 161.5, 138.4, 135.3, 135.1, 133.1, 130.3, 127.6, 125.0, 123.4, 120.2, 118.8, 111.9, 110.4, 100.0, 73.9, 46.6, 40.2, 28.4, 26.3, 24.0, 17.4; HRMS (ESI+) *m/z* calcd. for $C_{27}H_{28}N_4O_4$ [M+Na]$^+$ 495.20; found: 495.2021.

***tert*-Butyl(3-(7-acetyl-9-(3,4-dicyanophenyl)-3-oxo-1,3-dihydro-2H-pyrrol o[3,4-b]indolizin-2-yl)propyl)carbamate (19)**: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.67 (d, *J* = 9.0 Hz, 1H), 8.39 (s, 1H), 7.93 (s, 3H), 7.49 (d, *J* = 6.0 Hz, 1H), 5.25 (br s, 1H), 4.61 (s, 2H), 3.72 (t, *J* = 6.0 Hz, 2H), 3.19 (dd, *J* = 12.0, 6.0 Hz, 2H), 2.67 (s, 3H), 1.93-1.85 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.2, 161.2, 156.2, 139.5, 135.7, 135.4, 134.4, 131.4, 131.2, 130.9, 125.4, 123.9, 119.4, 117.0, 115.5, 115.4, 112.4, 111.0, 109.6, 79.3, 46.7, 40.3, 37.4, 29.7, 28.4, 26.4; LRMS (ESI+) *m/z* calcd. for $C_{28}H_{27}N_5O_4$ [M+H]$^+$ 498.21; found: 497.98.

***tert*-Butyl(3-(7-acetyl-3-oxo-9-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (20)**: Isolation yield was 39%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.57 (d, *J* = 6.0 Hz, 1H), 8.41 (s, 1H), 7.96 (d, *J* = 9.0 Hz, 2H), 7.56 (d, *J* = 6.0 Hz,

2H), 7.31 (d, *J* = 6.0 Hz, 1H), 5.45 (br s, 1H), 4.52 (s, 2H), 3.70 (t, *J* = 7.5 Hz, 2H), 3.18 (d, *J* = 6.0 Hz, 2H), 2.61 (s, 3H), 1.88-1.82 (m, 2H), 1.44 (s, 9H), 1.39 (s, 12H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.6, 161.9, 156.2, 136.2, 135.8, 135.4, 135.0, 134.8, 129.5, 127.9, 126.7, 124.8, 122.8, 121.3, 114.0, 109.7, 84.0, 75.1, 46.6, 40.2, 37.3, 31.9, 29.7, 28.4, 26.0, 24.9, 24.6; HRMS (ESI+) *m/z* calcd. for $C_{32}H_{40}BN_3O_6$ $[M+Na]^+$ 596.29; found: 596.2932.

tert-Butyl(3-(7-acetyl-9-(4-nitrophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (21): Isolation yield was 89%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.66 (d, *J* = 9.0 Hz, 1H), 8.45 (s, 2H), 8.39 (d, *J* = 9.0 Hz, 2H), 7.72 (d, *J* = 6.0 Hz, 2H), 7.39 (d, *J* = 9.0 Hz, 1H), 5.26 (br s, 1H), 4.59 (s, 2H), 3.73 (t, *J* = 6.0 Hz, 2H), 3.20 (dd, *J* = 12.0, 6.0 Hz, 2H), 2.67 (s, 3H), 1.93-1.83 (m, 2H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 195.3, 161.5, 156.1, 152.7, 146.0, 143.7, 140.5, 135.6, 135.3, 130.6, 127.5, 125.2, 124.8, 123.5, 120.2, 110.6, 79.2, 46.7, 44.7, 28.8, 28.4, 26.3; LRMS (ESI+) *m/z* calcd. for $C_{26}H_{28}N_4O_6$ $[M+Na]^+$ 515.19; found: 515.2.

tert-Butyl (3-(7-acetyl-9-(3,5-dimethoxyphenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (22): Isolation yield was 10%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.57 (d, *J* = 5.0 Hz, 1H), 8.46 (s, 1H), 7.30 (d, *J* = 10.0 Hz, 1H), 6.84 (d, *J* = 10.0 Hz, 2H), 6.49 (s, 1H), 4.52 (s, 2H), 3.69 (br s, 2H), 3.62 (br s, 6H), 3.13 (br s, 2H), 2.60 (s, 3H), 1.83-1.75 (m, 2H), 1.42 (s, 9H); LRMS (ESI+) *m/z* calcd. for $C_{28}H_{33}N_3O_6$ $[M+Na]^+$ 530.23; found: 530.3.

tert-Butyl (3-(7-acetyl-3-oxo-9-(thiophen-2-yl)-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (23): Isolation yield was 30%. ^{1}H

NMR (300 MHz, $CDCl_3$) δ 8.56-8.52 (m, 2H), 7.37-7.31 (m, 2H), 7.23-7.18 (m, 2H), 5.34 (br s, 1H), 4.52 (s, 2H), 3.71 (t, J = 6.0 Hz, 2H), 3.20 (d, J = 6.0 Hz, 2H), 2.67 (s, 3H), 1.94-1.82 (m, 2H), 1.45 (s, 9H); LRMS (ESI+) *m/z* calcd. for $C_{24}H_{27}N_3O_4S$ $[M+Na]^+$ 476.16; found: 476.3.

tert-Butyl 5-(7-acetyl-2-(3-((tert-butoxycarbonyl)amino)propyl)-3-oxo-2,3-dihydro-1H-pyrrolo[3,4-b]indolizin-9-yl)-1H-indole-1-carboxylate (24): Isolation yield was 15%. ^1H NMR (300 MHz, $CDCl_3$) δ 8.58 (d, J = 9.0 Hz, 1H), 8.44 (s, 1H), 8.29 (d, J = 9.0 Hz, 1H), 7.72 (s, 1H), 7.70 (d, J = 3.0 Hz, 1H), 7.50 (dd, J = 7.5, 1.5 Hz, 1H), 7.31 (s, 1H), 6.57 (d, J = 3.0 Hz, 1H), 5.40 (br s, 1H), 4.55 (s, 2H), 3.71 (t, J = 6.0 Hz, 2H), 3.21-3.17 (m, 2H), 2.62 (s, 3H), 1.90-1.84 (m, 2H), 1.73 (s, 9H), 1.44 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 198.3, 156.2, 152.7, 135.5, 134.7, 131.4, 129.1, 127.9, 126.9, 124.6, 124.1, 121.5, 119.9, 116.0, 114.8, 109.5, 107.3, 84,1, 40.1, 29.7, 28.4, 28.2, 26.1; LRMS (ESI+) *m/z* calcd. for $C_{33}H_{38}N_4O_6$ $[M+Na]^+$ 609.27; found: 609.3.

tert-Butyl (3-(7-cyano-9-(4-cyanophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (25): Isolation yield was 61%. ^1H NMR (400 MHz, $CDCl_3$) δ 8.60 (d, J = 8.0 Hz, 1H), 8.13 (s, 1H), 7.75 (d, J = 8.0 Hz, 2H), 7.58 (d, J = 8.0 Hz, 2H), 6.86 (d, J = 8.0 Hz, 1H), 5.20 (br s, 1H), 4.53 (s, 2H), 3.66 (t, J = 8.0 Hz, 2H), 3.14 (d, J = 4.0 Hz, 2H), 1.84-1.81 (m, 2H), 1.38 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.1, 156.0, 137.6, 135.5, 133.9, 133.1, 127.7, 125.8, 125.0, 124.1, 118.6, 117.9, 112.0, 111.5, 110.4, 104.7, 79.2, 46.5, 40.3, 37.3, 28.7, 28.4; LRMS (ESI+) *m/z* calcd. for $C_{26}H_{25}N_5O_3$ $[M+Na]^+$ 478.19; found: 478.2.

Methyl 2-(3-((*tert*-butoxycarbonyl)amino)propyl)-9-(4-cyanophenyl)-3-oxo-2,3-dihydro-1H-pyrrolo[3,4-b]indolizine-7-carboxylate (26): Isolation

yield was 71%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.54 (d, J = 8.0 Hz, 1H), 8.46 (s, 1H), 7.72 (d, J = 8.0 Hz, 2H), 7.60 (d, J = 8.0 Hz, 2H), 7.30 (dd, J = 6.0, 2.0 Hz, 1H), 5.29 (br s, 1H), 4.49 (s, 2H), 3.92 (s, 3H), 3.65 (t, J = 8.0 Hz, 2H), 3.14 (d, J = 4.0 Hz,2H), 1.83-1.81 (m, 2H), 1.38 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 165.5, 161.5, 156.0, 138.4, 135.2, 135.1, 133.0, 127.5, 124.8, 123.5, 123.1, 121.3, 118.8, 111.4, 111.1, 109.7, 52.6, 46.6, 40.1, 37.3, 28.7, 28.4; LRMS (ESI+) *m/z* calcd. for $C_{27}H_{28}N_4O_5$ $[M+Na]^+$ 511.20; found: 511.2.

tert-Butyl(3-(9-(4-cyanophenyl)-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (27): Isolation yield was 24%. ^{1}H NMR (400MHz, $CDCl_3$) δ 8.64 (d, J = 8.0 Hz, 1H), 7.93 (d, J = 12.0 Hz, 1H), 7.72 (d, J = 8.0 Hz, 2H), 7.61 (d, J = 8.0 Hz, 2H), 7.11 (t, J = 8.0 Hz, 1H), 6.83 (t, J = 9.0 Hz, 1H), 5.37 (br s, 1H), 4.51 (s, 2H), 3.69 (t, J = 12.0 Hz, 2H),3.18 (d, J = 4.0 Hz, 2H), 1.88-1.81 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.2, 156.2, 139.6, 136.8, 135.1, 133.0, 127.1, 125.9, 123.0, 121.3, 119.3, 118.1, 112.7, 108.8, 107.3, 79.2, 46.8, 40.1, 37.4, 29.8, 28.9, 28.5; LRMS (ESI+) *m/z* calcd. for $C_{25}H_{26}N_4O_3$ $[M+Na]^+$ 453.19; found: 453.2.

tert-Butyl (3-(9-(4-cyanophenyl)-7-methyl-3-oxo-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (28): Isolation yield was 28%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.52 (d, J = 8.0 Hz, 1H), 7.71 (d, J = 8.0 Hz, 2H), 7.60-7.58 (m, 3H), 6.67 (d, J = 8.0 Hz, 1H), 5.37 (br s, 1H), 4.48 (s, 2H), 3.67 (t, J = 8.0 Hz, 2H), 3.17 (d, J = 4.0 Hz, 2H), 2.92 (s, 3H), 1.87-1.81 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.1, 565.1, 139.7, 137.1, 135.4, 134.0, 132.8, 126.7, 125.2, 120.1, 119.2, 116.3, 115.1, 108.2, 105.8, 79.2, 46.7, 39.9, 37.2, 29.7, 28.4, 21.7, 14.1; LRMS (ESI+) *m/z* calcd. for $C_{26}H_{28}N_4O_3$ $[M+Na]^+$ 467.21; found: 467.2.

tert-Butyl (3-(7-cyano-3-oxo-9-phenyl-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (29): Isolation yield was 51%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.57 (d, J = 8.0 Hz, 1H), 8.12 (s, 1H), 7.49 (s, 4H), 7.39-7.36 (m, 1H), 6.79 (d, J =8.0 Hz, 1H), 5.31 (br s, 1H), 4.53 (s, 2H), 3.69 (t, J = 6.0 Hz, 2H), 3.17 (d, J = 4.0 Hz, 2H), 1.86-1.83 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.6, 156.1, 135.2, 133.7, 132.8, 129.5, 127.6, 127.5, 125.8, 125.4, 123.5, 118.4, 113.8, 111.3, 103.3, 79.2, 46.6, 40.3, 37.4, 28.8, 28.5; LRMS (ESI+) *m/z* calcd. for $C_{25}H_{26}N_4O_3$ $[M+Na]^+$ 453.19; found: 453.2.

Methyl 2-(3-((*tert*-butoxycarbonyl)amino)propyl)-3-oxo-9-phenyl-2,3-dihydro-1H-pyrrolo[3,4-b]indolizine-7-carboxylate (30): Isolation yield was 58%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.56 (d, J = 8.0 Hz, 2H), 7.56-7.48 (m, 4H), 7.34 (t, J = 8.0 Hz, 1H), 7.30 (d, J = 4.0 Hz, 1H), 5.39 (br s, 1H), 4.51 (s, 2H), 3.94 (s, 3H), 3.69 (t, J = 6.0 Hz, 2H), 3.18 (d, J = 4.0 Hz, 2H), 1.85-1.79 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 166.0, 162.1, 156.2, 134.9, 133.6, 129.4, 127.7, 126.9, 124.6, 122.5, 122.3, 122.2, 113.5, 110.9, 79.2, 52.5, 46.7, 40.2, 37.4, 28.8, 28.5; LRMS (ESI+) *m/z* calcd. for $C_{26}H_{29}N_3O_5$ $[M+Na]^+$ 486.20; found: 486.2.

tert-Butyl (3-(3-oxo-9-phenyl-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (31): Isolation yield was 14%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.56 (d, J = 8.0 Hz, 1H), 7.89 (d, J = 12.0 Hz, 1H), 7.51 (d, J = 8.0 Hz, 2H), 7.43 (t, J = 8.0 Hz, 2H), 7.26 (d, J = 8.0 Hz, 1H), 6.96 (t, J = 8.0 Hz, 1H), 6.72 (t, J = 9.0 Hz, 1H), 5.44 (br s, 1H), 4.45 (s, 2H), 3.65 (t, J = 9.0 Hz, 2H), 3.15 (dd, J = 12.0, 8.0 Hz, 2H), 1.83-1.75 (m, 2H), 1.40 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.4, 156.1, 136.2, 134.6, 134.5, 129.0, 127.1, 125.9, 125.4, 121.4, 120.1, 118.4, 111.8, 108.9,

78.9, 46.6, 39.8, 37.2, 28.8, 28.4; LRMS (ESI+) *m/z* calcd. for $C_{24}H_{27}N_3O_3$ $[M+Na]^+$ 428.20; found: 428.2.

***tert*-Butyl (3-(7-methyl-3-oxo-9-phenyl-1,3-dihydro-2H-pyrrolo[3,4-b]indolizin-2-yl)propyl)carbamate (32)**: Isolation yield was 29%. ^1H NMR (400 MHz, $CDCl_3$) δ 8.47 (d, J = 8.0 Hz, 1H), 7.57 (s, 1H), 7.52 (d, J = 8.0 Hz, 2H), 7.45 (t, J = 8.0 Hz, 2H), 7.28-7.24 (m, 1H), 6.59 (d, J = 8.0 Hz, 2H), 5.50 (br s, 1H), 4.45 (s, 2H), 3.66 (t, J = 6.0 Hz, 2H), 3.17 (dd, J = 12.0, 6.0 Hz, 2H), 2.37 (s, 3H), 1.83-1.77 (m, 2H), 1.43 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 162.5, 156.1, 136.7, 135.2, 134.7, 132.2, 129.0, 127.0, 125.7, 124.8, 119.5, 116.6, 114.5, 107.6, 79.0, 46.6, 39.8, 37.2, 28.8, 28.4, 21.6; LRMS (ESI+) *m/z* calcd. for $C_{25}H_{29}N_3O_3$ $[M+H]^+$ 442.21; found: 442.2.

***tert*-Butyl(3-(11-(4-(diethylamino)phenyl)-8-oxo-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (33)**: Isolation yield was 14%. ^1H NMR (300 MHz, $CDCl_3$) δ 8.40 (d, J = 9.0 Hz, 1H), 8.15 (d, J = 9.0 Hz, 1H), 7.62 (d, J = 9.0 Hz,1H), 7.42-7.35 (m, 2H), 6.92 (d, J = 9.0 Hz, 1H), 6.81 (d, J = 9.0 Hz, 1H), 5.52 (br s, 1H), 4.30 (s, 2H), 3.67 (t, J = 6.0 Hz, 2H), 3.17 (d, J = 6,0 Hz, 2H), 1.85-1.76 (m, 2H), 1.45 (s, 9H), 1.28 (s, 6H); LRMS (ESI+) *m/z* calcd. for $C_{32}H_{38}N_4O_3$ $[M+H]^+$ 527.30; found: 527.3.

***tert*-Butyl(3-(11-(4-methoxyphenyl)-8-oxo-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (34)**: Isolation yield was 54%. ^1H NMR (300 MHz, $CDCl_3$) δ 8.41 (d, J = 6.0 Hz, 1H), 7.95 (d, J = 6.0 Hz, 1H), 7.63 (d, J = 6.0 Hz, 1H), 7.47-7.38 (m, 3H), 7.30-7.27 (m, 1H), 7.05 (d, J = 6.0 Hz, 2H), 6.94 (d, J = 6.0 Hz, 1H), 5.49 (br s, 1H), 4.28 (s, 2H), 3.93 (s, 3H), 3.67 (t, J = 6.0 Hz, 2H), 3.19 (d, J = 12.0,

6.0Hz, 2H), 1.85-1.76 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 159.0, 152.6, 135.6, 131.6, 130.8, 128.3, 127.6, 127.3, 121.7, 126.7, 126.5, 123.2, 123.0, 121.5, 114.4, 112.0, 55.4, 46.0, 39.9, 32.4, 28.9, 28.5; LRMS (ESI+) *m/z* calcd. for $C_{29}H_{31}N_3O_4$ $[M+H]^+$ 486.24; found: 486.2.

tert-Butyl (3-(8-oxo-11-(p-tolyl)-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo [2,1-a]isoquinolin-9-yl)propyl)carbamate (35): Isolation yield was 40%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.41 (d, J = 6.0 Hz, 1H), 8.00 (d, J = 9.0 Hz, 1H), 7.63 (d, J = 6.0 Hz, 1H), 7.45-7.40 (m, 3H), 7.33-7.27 (m, 3H), 6.94 (d, J = 6.0 Hz, 1H), 5.49 (br s, 1H), 4.29 (s, 2H), 3.67 (t, J = 6.0 Hz, 2H), 3.18 (d, J = 6.0 Hz, 2H), 2.48 (s, 3H), 1.83-1.77 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.7, 156.1, 137.1, 135.7, 135.6, 132.5, 131.6, 129.7, 129.5, 128.3, 127.3, 127.1, 126.8, 126.4, 123.3, 123.0, 113.4, 112.1, 46.0, 39.9, 31,8, 28.8, 28.5, 21.4; LRMS (ESI+) *m/z* calcd. for $C_{29}H_{31}N_3O_3$ $[M+H]^+$ 470.24; found: 470.3.

tert-Butyl (3-(8-oxo-11-phenyl-8,10-dihydro-9H-pyrrolo[3',4' :4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (36): Isolation yield was 65%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.43 (d, J = 9.0 Hz, 1H), 7.97 (d, J = 9.0 Hz, 1H), 7.64 (d, J = 6.0 Hz, 1H), 7.57-7.39 (m, 6H), 7.29-7.24 (m, 1H), 6.96 (d, J = 1H), 5.49 (br s, 1H), 4.30 (s, 2H), 3.67 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 1.85-1.77 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.6, 156.1, 135.7, 135.6, 131.6, 129.6, 129.0, 128.4, 127.40, 127.37, 127.1, 126.9, 126.3, 123.2, 123.0, 113.4, 112.2, 104.1, 49.1, 39.9, 37.2, 29.7, 28.5, 14.2; LRMS (ESI+) *m/z* calcd. for $C_{28}H_{29}N_3O_3$ $[M+H]^+$ 456.23; found: 456.2.

tert-Butyl (3-(11-(4-fluorophenyl)-8-oxo-8,10-dihydro-9H-pyrrolo[3',4' :4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (37):I solation yield

was 56%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.42 (d, J = 6.0 Hz, 1H), 7.87 (d, J = 9.0 Hz, 1H), 7.65 (d, J = 9.0 Hz,1H), 7.53-7.48 (m, 2H), 7.42 (t, J = 6.0 Hz, 1H), 7.24-7.18 (m, 2H), 6.96 (d, J = 9.0 Hz, 1H), 5.46 (br s, 1H), 4.28 (s, 2H), 3.67 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 1.86-1.79 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.6, 156.1, 135.6, 131.4, 131.3, 128.4, 127.5, 127.2, 126.9, 126.2, 123.0, 116.2, 115.9, 112.2, 79.0, 45.9, 39.9, 37.2, 29.7, 28.9, 28.5; LRMS (ESI+) *m/z* calcd. for $C_{28}H_{28}FN_3O_3$ $[M+H]^+$ 474.22; found: 474.2.

tert-Butyl (3-(8-oxo-11-(4-(trifluoromethyl)phenyl)-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (38): Isolation yield was 69%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.44 (d, J = 6.0 Hz, 1H), 7.92 (d, J = 6.0 Hz, 1H), 7.78-7.66 (m, 5H), 7.49-7.43 (m, 1H), 7.34-7.29 (m, 1H), 7.00 (d, J = 6.0 Hz, 2H), 5.43 (br s, 1H), 4.31 (s, 2H), 3.68 (t, J = 6.0 Hz, 2H), 3.19 (dd, J = 12.0, 6.0 Hz, 2H), 1.86-1.80 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 162.4, 156.1, 143.8, 139.6, 135.6, 131.7, 129.8, 128.5, 127.6, 127.4, 127.3, 126.0, 125.9, 123.1, 123.0, 122.2, 112.6, 79.1, 46.0, 40.0, 37.2, 28.9, 28.4; LRMS (ESI+) *m/z* calcd. for $C_{29}H_{28}F_3N_3O_3$ $[M+Na]^+$ 546.20; found: 546.1.

tert-Butyl (3-(11-(4-cyanophenyl)-8-oxo-8,10-dihydro-9H-pyrrolo[3',4':4,5]pyrrolo[2,1-a]isoquinolin-9-yl)propyl)carbamate (39): Isolation yield was 55%. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.45 (d, J = 9.0 Hz, 1H), 8.38 (d, J = 6.0 Hz, 1H), 8.10 (d, J = 9.0 Hz, 1H), 7.91 (d, J = 6.0 Hz, 1H), 7.80 (d, J = 9.0 Hz, 1H), 7.70-7.65 (m, 2H), 7.57-7.45 (m, 2H), 6.95-6.90 (m, 2H), 5.51 (br s, 1H), 4.36 (s, 2H), 3.68 (t, J = 6.0 Hz, 2H), 3.19 (d, J = 3.19 Hz, 2H), 1.88-1.81 (m, 2H), 1.46 (s, 9H); LRMS (ESI+) *m/z* calcd. for $C_{29}H_{28}N_4O_3$ $[M+Na]^+$ 503.21; found: 503.2.

2.4.3 Cell Culture and Imaging

HeLa cells were seeded on cover glass-bottom petri dish and incubated at 5 % CO_2 incubator at 37 ° C overnight. HeLa cells were incubated with a medium containing compound **20** (80 μM) for 10 min. After PBS washing, fluorescent signals of the treated cells were measured by using fluorescence microscopy. Then, the HeLa cells treated with compound **20** were exposed to 1% H_2O_2 solution diluted in PBS. For the consistency of bioimaging, we maintained the identical view of HeLa cells pre- and post-treated with H_2O_2 solution. The fluorescence images were captured by the identical exposure time and PMT voltage to maintain the consistency of this imaging experiment.

2.5 References

[1] a) J. R. Lakowicz, Principles of Fluorescence Spectroscopy, Springer, New York, **2006**; b) W. Rettig, B. Strehmel, S.Schrader, H. Seifert, Applied Fluorescence in Chemistry, Biology, and Medicine, Springer, New York, **1999**; c) R. P. Haugland, Handbook of Fluorescence Probes and Research Products, Molecular Probes, Eugene, OR, **2002**; d) J. J. Lavigne, E. V. Anslyn, *Angew. Chem.* **2011**, *123*, 3174-3175; *Angew. Chem. Int. Ed.* 2011, *50*, 3118-3130; e) J. Zhang, R. E. Campbell, A. Y. Ting, R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 906-918; f) D. W. Domaille, E. L. Que, C. J. Chang, *Nat. Chem. Biol.* **2008**, *4*, 168-175.

[2] A. P. Demchenko, Advanced Fluorescence Reporters in Chemistry and Biology I, Springer, Berlin, **2010**.

[3] a) T. Miura, Y. Urano, K. Tanaka, T. Nagano, K. Ohkubo, S. Fukuzumi, *J. Am. Chem. Soc.* **2003**, *125*, 8666-8671; b) T. Ueno, Y. Urano, K. Setsukinai, H. Takakusa, H. Kojima, K. Kikuchi, K. Ohkubo, S. Fukuzumi, T. Nagano, *J. Am. Chem. Soc.* **2004**, *126*, 14079-14085; c) Y. Urano, M. Kamiya, K. Kanda, T. Ueno, K.Hirose, T. Nagano, *J. Am. Chem. Soc.* **2005**, *127*, 4888-4894; d) T. Ueno, Y. Urano, H. Kojima, T. Nagano, *J. Am. Chem. Soc.* **2006**, *128*, 10640-10641.

[4] a) E. Kim, M. Koh, J. Ryu, S. B. Park, *J. Am. Chem. Soc.* **2008**, *130*, 12206-12207; b) E. Kim, M. Koh, B. J. Lim, S. B. Park, *J. Am. Chem. Soc.* **2011**, *133*, 6642-6649.

[5] a) E. Kim, S. Lee, S. B. Park, *Chem. Commun.* **2012**, *48*, 2331-2333; b) Y. Lee, S. Na, S. Lee, N. L. Jeon, S. B. Park, *Mol. BioSyst.* **2013**, *9*, 952-956; d) M. S. Jeong, E. Kim, H. J. Kang, E. J. Choi, A. R. Cho, S. J. Chung, S. B. Park, *Chem. Commun.* **2012**, *48*, 6553-6555.

[6] a) A. S. K. Hashmi, *Chem. Rev.* **2007**, *107*, 3180-3211; b) S. Su,J. A. Porco, Jr., *J. Am. Chem. Soc.* **2007**, *129*, 7744-7745.

[7] a) X. Chen, K. M. Engle, D. Wang, J. Yu, *Angew. Chem.* **2009**, *121*, 5196-5217; *Angew. Chem. Int. Ed.* **2009**, *48*, 5094-5115; b) J. Dupont, C. S. Consorti, J. Spencer, *Chem. Rev.* **2005**, *105*, 2527-2572; c) S. H. Cho, J. Y. Kim, J. Kwak, S. Chang, *Chem. Soc. Rev.* **2011**, *40*, 5068-5083; d) T. W. Lyons, M. S. Sanford, *Chem. Rev.* **2010**, *110*, 1147-1169.

[8] a) Z. Liang, J. Zhao, Y. Zhang, *J. Org. Chem.* **2010**, *75*, 170-177; b) D. R. Stuart, E. Villemure, K. Fagnou, *J. Am. Chem. Soc.* **2007**, *129*, 12072-12073; c) M. Miyasaka, K. Hirano, T. Satoh, M. Miura, *J. Org. Chem.* **2010**, *75*, 5421-5424.

[9] C. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165-195.

[10] T. Terai, R. Tomiyasu, T. Ota, T. Ueno, T. Komatsu, K. Hanaoka, Y. Urano, T. Nagano, *Chem. Commun.* **2013**, *49*, 3101-3103.

[11] a) A. P. de Silva, H. Q. N. Gunaratne, T. Gunnlaugsson, A. J. M. Huxley, C. P. McCoy, J. T. Rademacher, T. E. Rice, *Chem. Rev.* **1997**, *97*, 1515 - 1566; b) M. Kollmannsberger, K. Rurack, U.Resch-Genger, J. Daub, *J. Phys. Chem. A* **1998**, *102*, 10211-10220.

[12] a) M. C. Y. Chang, A. Pralle, E. Y. Isacoff, C. J. Chang, *J. Am. Chem. Soc.* **2004**, *126*, 15392-15393; b) B. C. Dickinson, C. J. Chang, *J. Am. Chem. Soc.* **2008**, *130*, 9638-9639; c) A. E. Albers, V. S. Okreglak, C. J. Chang, *J. Am. Chem. Soc.* **2006**, *128*, 9640-9641; d) M. Abo, Y. Urano, K. Hanaoka, T. Terai, T. Komatsu, T.Nagano, *J. Am. Chem. Soc.* **2011**, *133*, 10629-10637.

[13] R. E. DeSimone, R. S. Drago, *J. Am. Chem. Soc.* **1970**, *92*, 2343-2352.

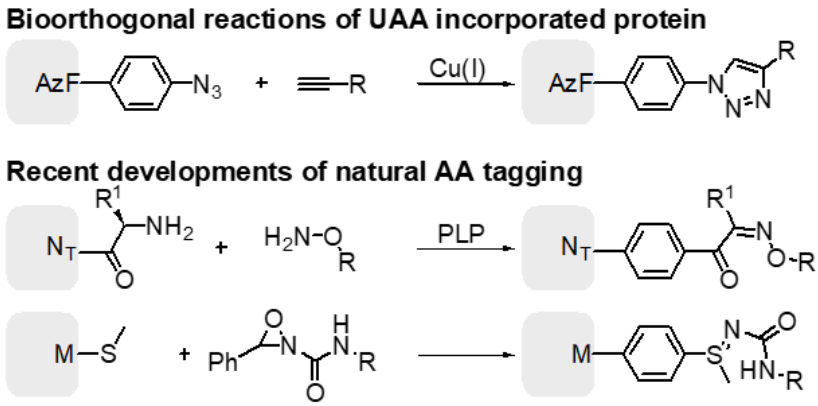
[14] W. Zhu, D. Ma, *Org. Lett.* **2006**, *8*, 261-263

Chapter 3. Discovery of Tyrosine Selective Bioconjugation

3.1 Introduction

Biological regulatory system works through control of protein functions. Many protein- or peptide-based drugs have been actively applied for regulating various signaling pathway, however, their instability in physiological condition results short duration of action, thus hampering wide practical applications. Therefore, artificial functional groups are often introduced for overcoming the disadvantages.

Site-specific functionalization of biomolecules is of surmounting importance in chemical biology studies, since non-selective functionalization generates heterogeneous mixtures with often decreased activities to unpredictable extents.[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 certain natural amino acid residue (Scheme 3.1).[4] As an example of the former, insertion of azidophenylalanine (AzF) has been carried out 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 and the expression yield is often moderate.[5]

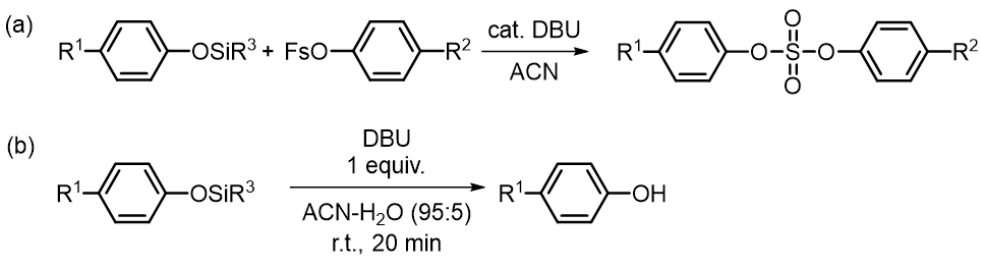


Scheme 3.1. Protein functionalization using the click reaction of genetically modified proteins with an UAA (first row) and recent discoveries of bioconjugations to natural amino acids in native proteins.

There are ample examples of readily accessible natural protein functionalization methods in the second category. Lys-succinimide and 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 are 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 a strategy for an *N*-terminus modification through pyridoxal phosphate (PLP) mediated transamination and subsequent oxime formation.[9] Toste and Chang's group discovered a selective Met bioconjugation using oxaziridine reagents.[10] Functionalization on only one *N*-terminus or a rare amino acid such as Met appears to be attractive for the maximization of selectivity;[11] 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.

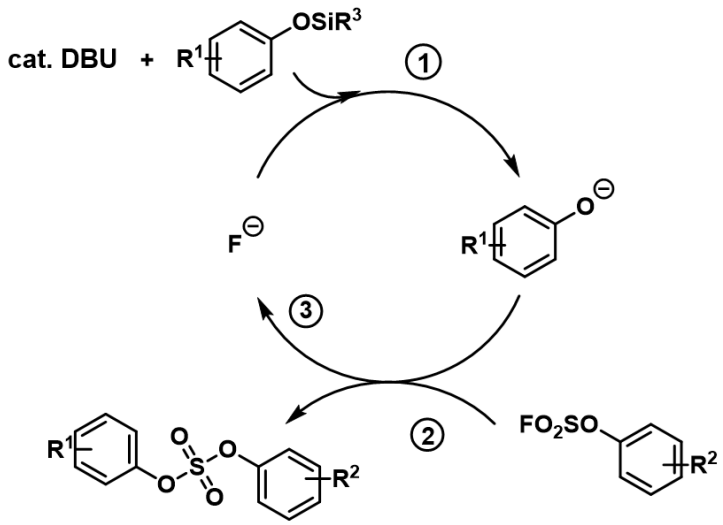
3.2 Results and Discussion

3.2.1 SuFEx Reaction of Aryloxy Anion and Aryl Fluorosulfate



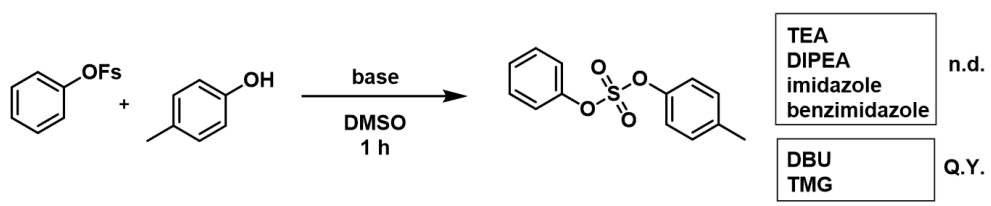
Scheme 3.2. Comparison between SuFEx reaction developed by Sharpless' group and stoichiometric DBU-mediated desilylation reaction.

Recently, Sharpless' group developed a sulfate click reaction between an aryl silyl ether and an aryl fluorosulfate (fosylate) in the presence of a *catalytic amount* of a strong base such as DBU, named SuFEx reaction (Scheme 3.2a).[12] Though extremely useful, this reaction has not been widely utilized for Tyr bioconjugation since silylated Tyr residue was presumed to be required.[13] In previous work of our research group, desilylation of various aryl silyl ether with a stoichiometric amount of DBU in polar environment was reported (Scheme 3.2b).[14] Therefore we hypothesized that in the SuFEx reaction the initial deprotection of an aryl silyl ether mediated by a catalytic amount of DBU generates an aryl oxide anion, which in turn reacts with the fosylate producing a fluoride anion, hence the completion of the catalytic cycle (Scheme 3.3).



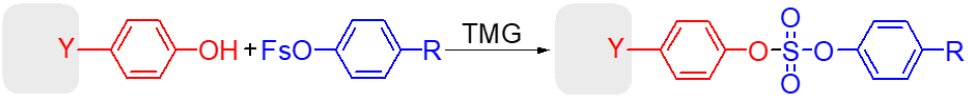
Scheme 3.3. A proposed mechanism for SuFEx reaction between aryl fosylate and aryl silyl ether.

This hypothesis obviates the need of an aryl silyl ether in our 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 (Scheme 3.4).



Scheme 3.4. Base screening in modified SuFEx reaction between *p*-cresol and phenyl fosylate. n.d. denotes non-detection.

With 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 (Scheme 3.5).

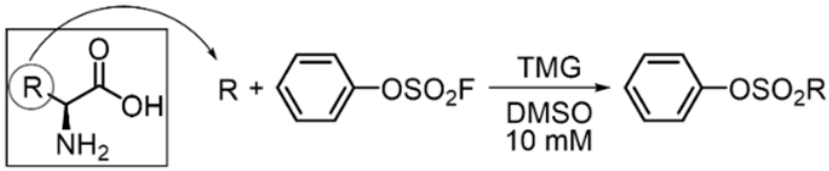


Scheme 3.5. Proposed selective Tyr modification in native proteins through the SuFEx chemistry.

There has been a few reports on Tyr-selective bioconjugation using Mannich or ene-type reactions.[15] However, those reactions 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 Tyr residues at a binding site of LBP (lipid binding protein) protein.[13] An Arg residue around the binding site appeared to assist the bond formation between the bioprobe and Tyr. This report was in accord with our experimental results, as the guanidine side chain of Arg is structurally similar to TMG. While the previous study reported the reaction between Tyr only at the binding site of the protein and fosylate probe,[13, 16] we envisioned that the application of SuFEx reaction could be expanded to universal Tyr functionalization on the surface of proteins.

3.2.2 Chemoselectivity of SuFEx Reaction among Various Nucleophilic Amino Acid Model Compounds

As the first step of developing the protein conjugation strategy, we studied the chemoselectivity of the modified SuFEx on a model-conjugating molecule, phenyl fosylate, towards various model counterpart molecules corresponding to representative nucleophilic amino acid residues (Table 3.1).



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

Table 3.1. Comparison of SuFEx reactivity in the reactions of various model nucleophiles with phenyl fosylate. [a] Isolation yield. [b] 0.25 equiv. of $NiCl_2(H_2O)_6$ was added.

Propanethiol was selected as a model nucleophile of Cys instead of gaseous methanethiol, for the ease of handling. 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*-tolylphenyl 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 sulfamide, or thiosulfate was generated even in 12 h, which implies that the reaction of the fosyl group with Lys or Cys (Entry 2, and 3) is unlikely. As expected, methanol did not participate in the reaction with fosylate (Entry 4). Although guanidine works as a base, it did not act as a nucleophile (Entry 5). Therefore, we expected that it would be highly

unlikely for Ser, Thr or Arg to participate in the SuFEx. we also tested the reactivity of the heterocycles embedded in amino acid residues: reaction with 3-methylindole yielded 12.5% of the sulfamide product after 12 h (Entry 6). This indicates that the SuFEx reaction with Trp may be too slow to compete the Tyr-SuFEx reaction. Though 4-methylimidazole, a nucleophile corresponding to His, was converted into the sulfamide (Scheme 3.6), adding a small amount of Ni^{2+}, which is the 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.

Scheme 3.6. Reaction between phenyl fosylate and imidazole to yield a sulfamide product.

3.2.3 Fluorescence Tagging on TAT 47-57 Peptide.

Then we assessed the possibility of the SuFEx application in the peptide labeling. We took TAT 47-57, the most well-known cell penetrating peptide (CPP),[17] 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.

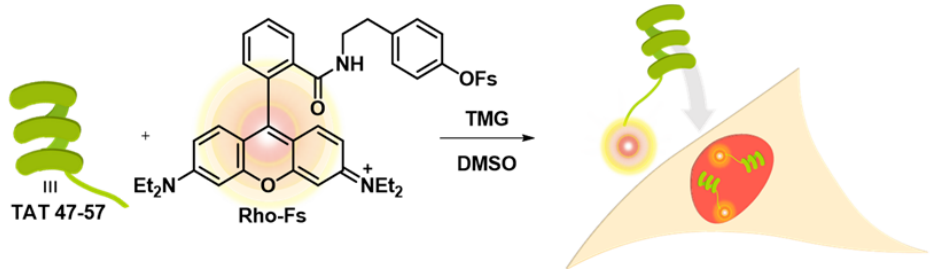


Figure 3.1. Functionalization of TAT 47-57 with a fluorescent small molecule, Rho-Fs.

TAT 47-57, upon reaction with Rho-Fs, yielded Rho-TAT quantitatively (Figure 3.1). The detected *m/z*, 2183.19, by MALDI-TOF MS, matched well the expected mass of Rho-TAT (Figure 3.2). No multi-substituted products were generated.

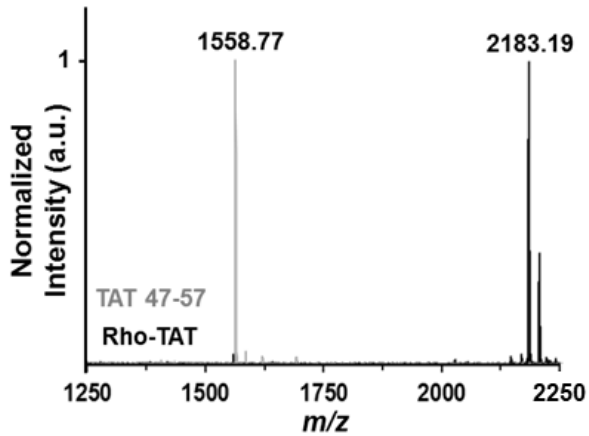


Figure 3.2. MALDI-TOF MS spectra of TAT 47-57 (*m/z* expected; 1558.97 / observed; 1558.77) and Rho-TAT (expected; 2183.23 / observed; 2183.19).

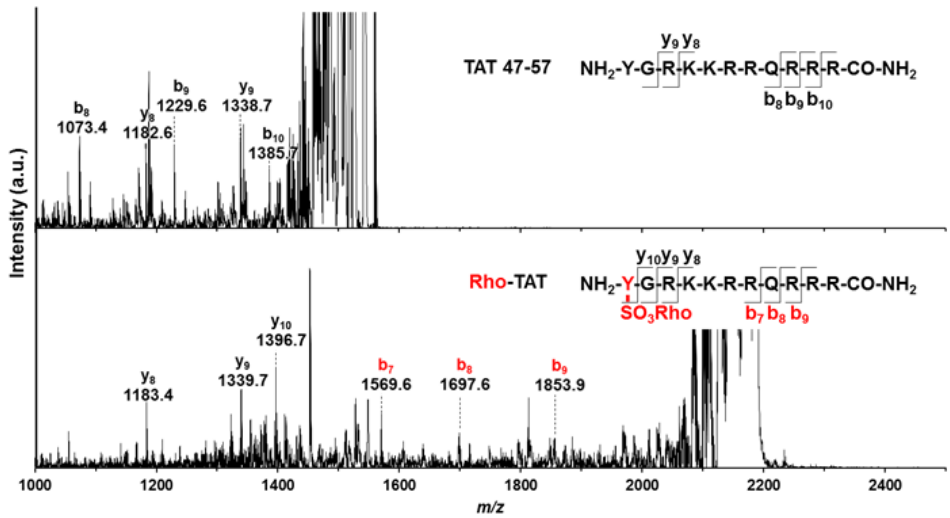


Figure 3.3. MALDI-TOF/TOF MS spectra of TAT 47-57 (upper) and Rho-TAT.

To ensure 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 3.3). 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 3.4).[18]

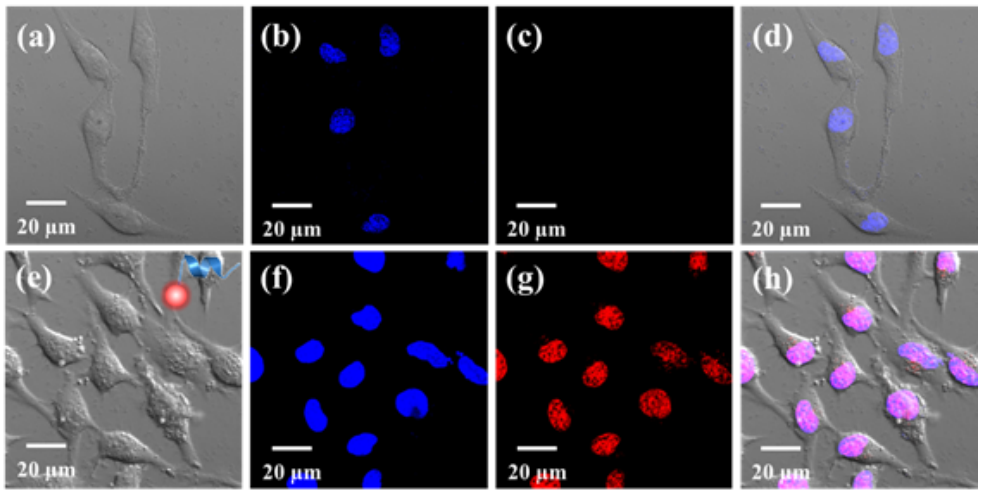


Figure 3.4. Confocal laser scanning microscopy images of HeLa cells after 12 h-treatment of PBS (a-d), and Rho-TAT peptide (e-h). Nucleus was stained with Hoechst 33342 (blue).

Accordingly, we could conclude 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.

3.2.4 PEGylation of Erythropoietin

After the successful functionalization of the short peptide, we proceeded further to protein functionalization in aqueous conditions. Recombinant human erythropoietin (rhEPO) was selected 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.[19] To improve the pharmacokinetic stability of rhEPO for less frequent administration, the conjugation of hydrophilic polymers such as poly(ethyleneglycol) (PEG) is widely used.[20] PEG-conjugation,

i.e. PEGylation, forms a hydrated polymeric layer around the proteins to reduce immunogenicity, renal clearance, and enzymatic degradation.[21] Mircera, a PEGylated form of rhEPO, shows a significantly prolonged half-life, allows much less frequent administration than that of unmodified ones. The PEGylation on Mircera is based on th e Lys-succinimide chemistry. Although the succinimide predominantly reacts with either Lys-45 or Lys-52 among seven Lys residues on the surface of rhEPO,[22] a heterogeneous mixture of isomers is produced. The intrinsic activity of each isomer has not been documented because of the purification difficulty. However, 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.[23] Therefore, Tyr-49 appeared to be a suitable site for the SuFEx-mediated PEGylation without any interruption in protein function.

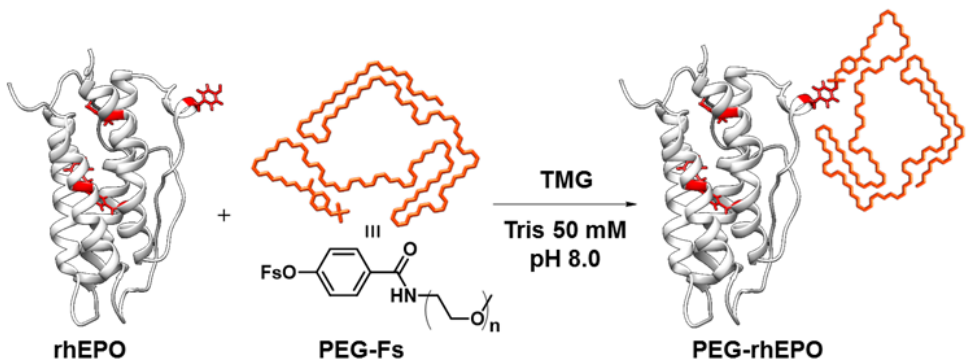


Figure 3.5. Schematic representation of the SuFEx reaction between rhEPO and a PEGylating reagent.

We synthesized a short fosylated PEG (PEG-Fs; M_n = 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 (Figure 3.5). 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 3.6).

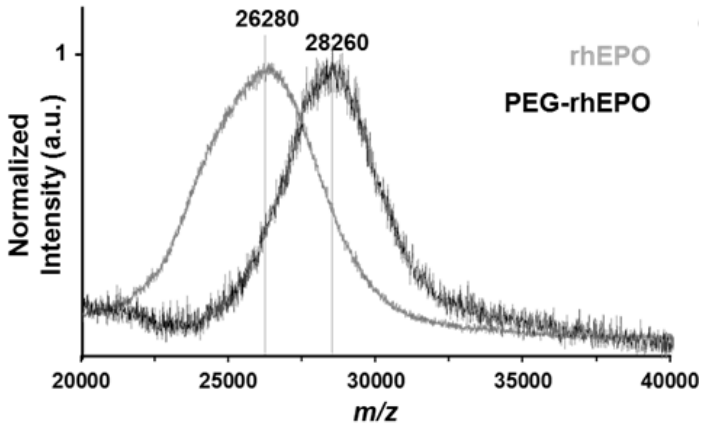


Figure 3.6. MALDI-TOF MS spectra of rhEPO (*m/z* observed; 26,280) and PEG-rhEPO (observed; 28,260).

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 the Tyr-49 (Figure 3.7, Table 3.2) and none of the internal Tyr residues underwent PEGylation. This result implies that the structural integrity of rhEPO 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.

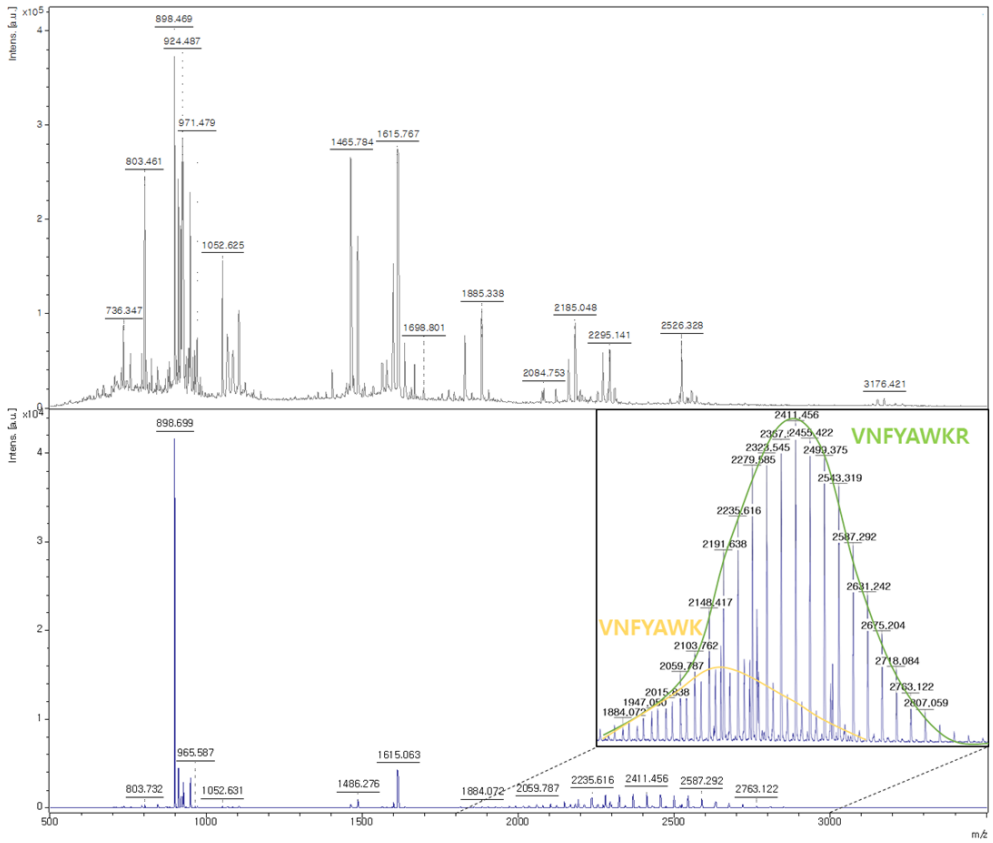


Figure 3.7. MALDI-TOF MS spectra of rhEPO (upper) and PEG-rhEPO after trypsin treatment.

VNFYAWKR (Y49)										
n (# of PEG)	15	16	17	18	19	20	21	22	23	24
Expected	1970.991	2015.018	2059.044	2103.07	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	2279.585	2323.545	2367.508
n (# of PEG)	25	26	27	28	29	30	31	32	33	34
Expected	2411.254	2455.28	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

VNFYAWK (Y49)							
n (# of PEG)	18	19	20	21	22	23	24
Expected	1946.969	1990.995	2035.021	2079.048	2123.074	2167.1	2211.126
Observed	1947.05	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	2342.61	2386.546	2430.476	2474.422	2518.322

Table 3.2. PEGylated *m/z* found in MALDI-TOF spectrum of PEG-rhEPO after trypsin treatment; mis-cleaved VNFYAWKR (upper) and fully cleaved VNFYAWK peptides.

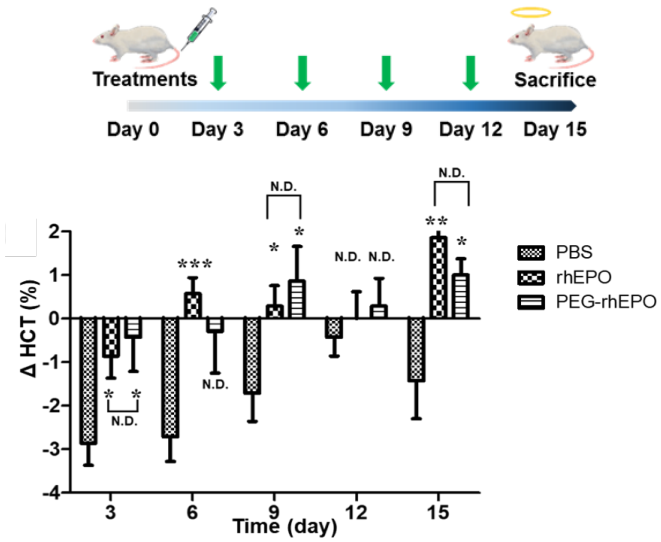


Figure 3.8. Schematic representation of mice experiment schedule (upper) and the HCT profiles of control, rhEPO-treated, PEG-rhEPO-treated Balb/c mice.

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 a dose of 20 μg/kg every three days. *In vivo* activities were compared by the measurement of the hematocrit (HCT).[24] The control group showed a significant decreases 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 3.8). 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 PEG-rhEPO-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.

3.3 Conclusion

A new chemoselective reaction for the Tyr functionalization is developed. We developed a modified SuFEx chemistry, where phenols can directly react with an aryl fosylate without the need of pre-suilylation. With the midified SuFEx chemistry, tyrosine was chemoselectively functionalized. It was confirmed that among various nucleophilic amino acid residues, only Tyr exclusively reacts with aryl fosylate.

The applicability of the SuFEx chemistry to the peptide and protein modifications was clearly demonstrated without disrupting their own functions. A tyrosine in cell-penetrating TAT 47-57 peptide was conjugated with rhodamine fosylate, and the resulting fluorescence-conjugated TAT peptide penetrated into HeLa cell nucleus, which implies that both the fluorescence and cell penetration functions were well preserved.

Furthermore, a surface-exposed tyrosine in rhEPO could be selectively PEGylated. Bioconjugation reaction of rhEPO underwent well in a buffered solution (Tris pH 8.0), which warrants the preservation of the original biomolecular structure. The singularly PEGylated rhEPO retained hematogenous funtion as confirmed by a mouse *in vivo* experiment.

As the reaction condition of the SuFEx chemistry is simple, chemoselective, and non-destructive, we anticipate a broad utility of this Tyr selective bioconjugation methods.

3.4 Experimental Section

3.4.1 Materials and Methods

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*-Dimethyl formaide (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 (1x) was purchased from Gibco. Erythropoietin (EPO; human) was purchased from GenScript. Centrifugal Filter Units were purchased from Milipore.

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, $CDCL_3$, or DMSO-D_6 for neutral compounds.

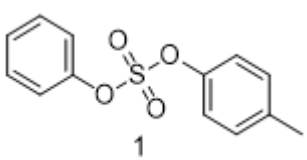
Reactions that needed anhydrous conditions were carried out in flame-dried glassware under positive pressure of dry Ar using standard Schlenk line techniques. Evaporation of solvents was performed at reduced pressure using a rotary evaporator. Reactions of small

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 NanoDrop™2000 spectrophotometer.

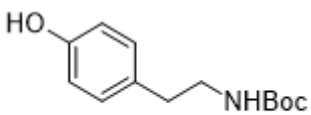
3.4.2 Preparations of Compounds, Peptides and Proteins



1

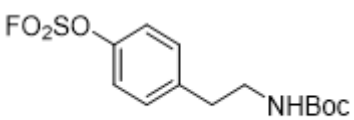
Phenyl *p*-tolyl sulfate (1): A mixture of *p*-cresol (21.6 mg, 0.200 mmol)

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.



2

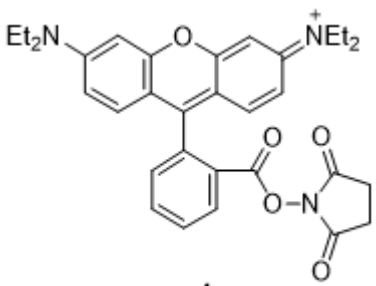
***N*-Boc tyramine (2):** *N*-Boc tyramine was synthesized through a procedure described in the literature.[25]



3

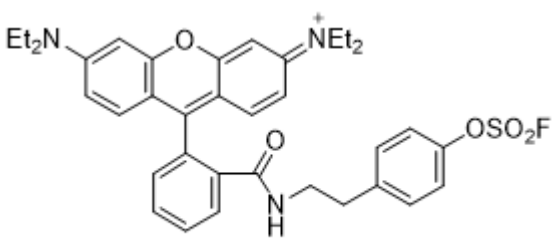
4-(2-((*tert*-Butoxycarbonyl)amino)ethyl)phenyl sulfurofluoridate (3): A mixture of compound **2** (0.130 g, 0.550 mmol) and triethylamine (0.240 mL, 1.69 mmol) was stirred in DCM (2.75 mL) under SO_2F_2 atmosphere at room temperature. After 5h, the mixture was poured into distilled water (20.0mL) and extracted with MC (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 (145mg, 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.



4

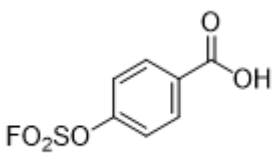
Tetraethylrhodamine succinimidyl ester (4): *N*-Boc tyramine was synthesized through a procedure described in the literature.[26]



5

***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 MC (30.0 mL) 3 times. The organic layer was dried over $MgSO_4$, filtered and concentrated under reduced pressure. To

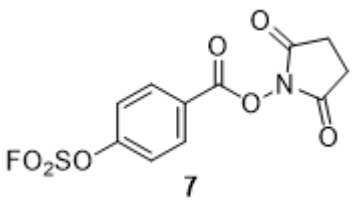
the residue in anhydrous DCM (2.00 mL) were added compound **4** (40.0 mg, 0.070 mmol) and TEA (20.0 μL, 0.140 mmol) and the mixture was stirred in for 2 h. The reaction mixture was then poured into brine (10.0 mL) and extracted with MC (10.0 mL) 3 times. The organic layer was dried over $MgSO_4$, and then was purified with column chromatography (EA/Hex). A yellow oil (11mg, 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 ($CDCl_3$, 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.



6

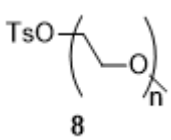
4-((Fuorosulfonyl)oxy)benzoic acid (6): A mixture of 4-hydroxy benzoic acid (276 mg, 2.00 mmol), and TEA (1.10 mL, 8.00 mmol) was stirred under sulfuryl fluoride atmosphere in DCM (5.00 mL) for 12 h. The mixture was poured into 1 N HCl solution (10.0 mL) and extracted with MC (10.0 mL) 3 times. The organic 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.



7

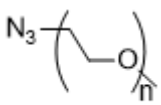
2,5-Dioxopyrrolidin-1-yl 4-((fluorosulfonyl)oxy)benzoate (7): A mixture of compound **6** (113 mg, 0.510 mmol), *N*-hydroxysuccinimide (226 mg, 2.00 mmol), and *N,N'*-dicyclohexylcarbodiimide (117 mg, 0.570 mmol) was stirred in anhydrous THF (5.00 mL) for 12 h. The mixture was concentrated under reduced 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.



8

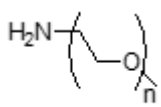
Methoxy poly ethylene glycol tosylate (8): Methoxy poly ethylene glycol-2000 (MPEG-2000, 2.30 g, 1.15 mmol), TsCl (0.400 g, 2.10 mmol), and TEA (0.400 mL, 2.90 mmol) were stirred in DCM (2.00 mL) at 50 °C for overnight. The mixture was poured into 1 N NaOH solution (40.0 mL) and extracted with MC (10.0 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. ^1H 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).



9

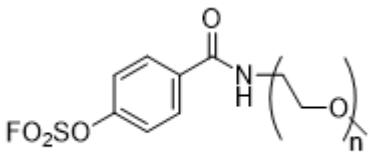
Methoxy poly ethylene glycol azide (9): A mixture of compound **8** (400 mg, 0.190 mmol), sodium azide (24.2 mg, 0.370 mmol), and sodium bicarbonate (23.5 mg, 0.280 mmol) was stirred in DMF (1.0 mL) at 120 °C for 12 h. The solvent was completely evaporated, and the residue dissolved in a small amount of DCM and triturated with ethyl ether (5.0 mL). A yellow solid (363 mg, 96.4%) was obtained.



10

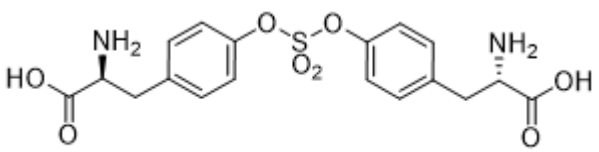
Methoxy poly ethylene glycol amine (10): A mixture of compound **9** (363 mg, 0.180 mmol) and triphenylphosphine (70.5 mg, 0.270 mmol) was stirred in anhydrous THF (1.00 mL) for 12 h. Water (100 μL) was added to the mixture and it was stirred for 12 h. The solvent was evaporated, and the residue was 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).



11

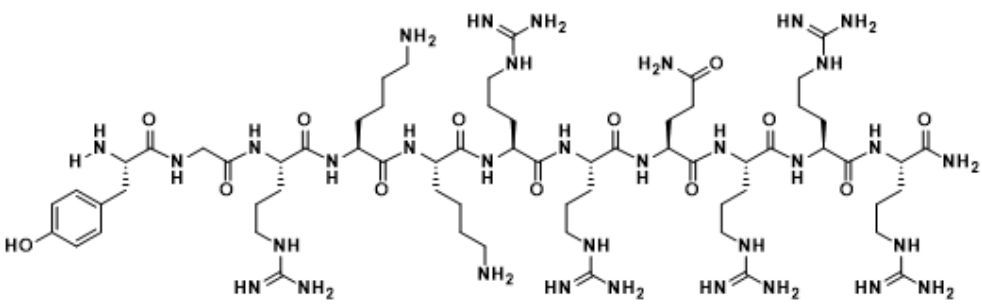
4-(Methoxy poly ethylene glycolyl carbamoyl)phenyl sulfurofluoridate (11): A mixture of compound **7** (17.5 mg, 0.055 mmol), **10** (100 mg, 0.05 mmol), and TEA (14 μL, 0.1 mmol) was stirred in anhydrous THF. After 12 h, the mixture was poured into 1 N HCl solution and extracted with DCM. The organic 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.



12

(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. ^1H 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.

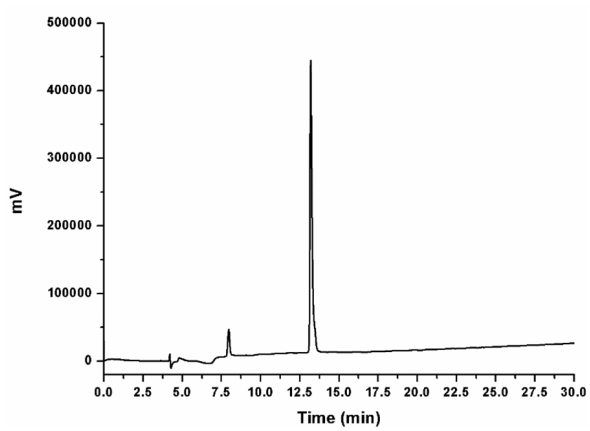


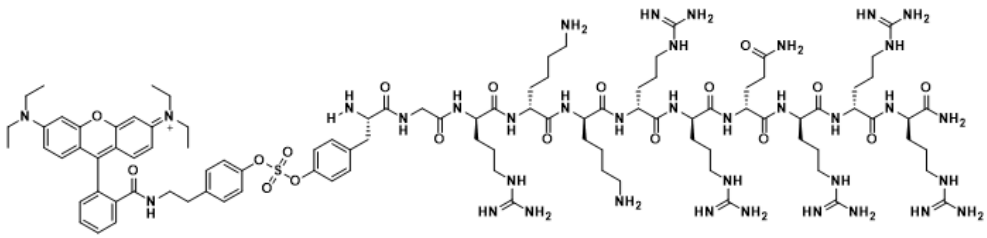
TAT 47-57 peptide: Synthesis of TAT peptide (sequence: YGRKKRRQRRR) was performed by using microwave heating in Fmoc-based solid-phase peptide synthesis with Rink Amide MBHA resin. Peptides were synthesized in 90 μmol scale. The typical conditions[27] for microwave-assisted peptide synthesis were used: Fmoc-deprotection with 20% piperidine at 75 °C (50 W) for 20 min and peptide coupling reaction with PyBOP and DIPEA at 75 °C (25 W) for 20 min.

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 Fmoc-Arg(Pbf)-OH (350.3 mg, 0.540 mmol) were stirred with

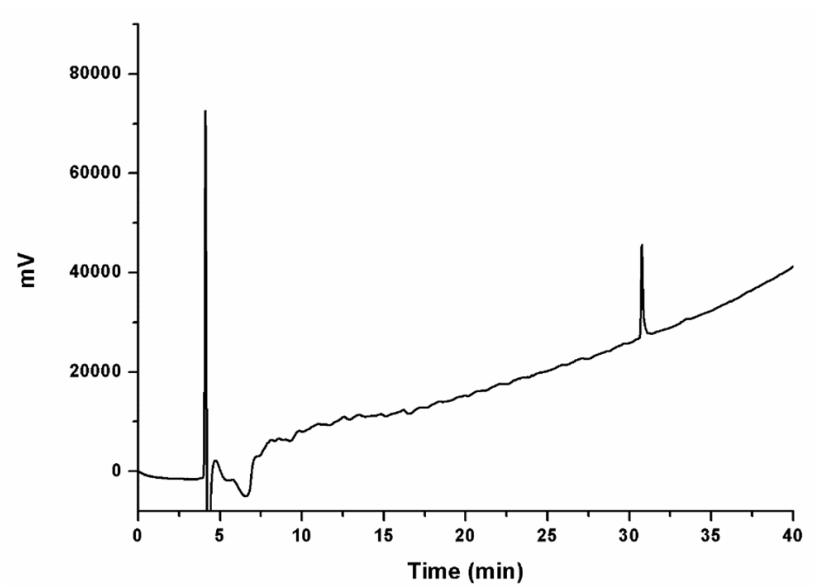
PyBOP (281 mg, 0.540 mmol), and DIPEA (200 μL, 1.18 mmol) under N_2 atmosphere. Byproducts were removed with DMF and DCM several times. After the coupling step, Fmoc-protected amino acid was deprotected with 20% piperidine in DMF, and the byproducts were removed with DMF and DCM several times. The coupling and deprotection steps were repeated with different Fmoc-protected amino acids sequentially until the deprotection of the last Fmoc-protected amino acid.

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 is shown below (>93% purity, absorbance was measured at 220 nm).





Rho-TAT peptide: The purified TAT 47-57 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 is shown below (>99% purity, absorbance was measured at 220 nm).



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 (x5). Then, the buffer was exchanged into DPBS by using Amicon Ultra-0.5 devices (x3). LRMS (MALDI) *m/z*: Anal. calcd. for $[M+H]^+$: 28914.9; found: 28260.2.

Tryptic digestion of proteins: rhEPO and PEG-rhEPO were subjected to trypsin digestion through a modified method.[2x] 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.

3.4.3 Cell Culture and Imaging

HeLa cells (10,000 cells/well) were cultured in 180 μL of DMEM supplemented with 10% FBS on confocal dishes at 37 °C in a 5% CO_2 incubator. After 24 h, the media was exchanged with fresh DMEM supplemented with 2% FBS. Then, Rho-TAT peptide in PBS (final concentration: 10 μL) were added to the dish and the cells were further incubated for 12 h. Cells were washed 3 times with fresh media for the removal of peptides outside the cells. Nucleus was

stained with a Hoechst 33342 solution (final concentration: 4 μL) for 10 min prior to the confocal laser scanning microscopy imaging. After further washing with fresh media (x3), 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.

3.4.4 *In Vivo* Mice Experiment

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.

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 (△ HCT) was determined by the difference between the initial hematocrit (day 0) and the hematocrit at each time point for each mouse.

3.5 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**, *8*, 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] a) J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi, M. B. Francis, *Angew. Chem. Int. Ed.* **2006**, *45*, 5407-5311; b) 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.

[10] 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.

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

[12] J. Dong, L. Krasnova, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **2014**, *53*, 9430-9448.

[13] W. Chen, J. Dong, L. Plate, D. E. Mortenson, G. J. Brighty, S. Li, Y. Liu, A. Galmozzi, P. S. Lee, J. J. Hulce, *J. Am. Chem. Soc.* **2016**, *138*, 7353-7364.

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

[15] a) H. Ban, J. Gavrilyuk, C. F. Barbas III, *J. Am. Chem. Soc.* **2010**, *132*, 1523-1525; b) N. S. Joshi, L. R. Whitaker, M. B. Francis, *J. Am. Chem. Soc.* **2004**, *126*, 15942-15943.

[16] 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.

[17] a) H. Brooks, B. Lebleu, E. Vives, *Adv. Drug. Deliv. Rev.* **2005**, *57*, 559-577; b) X. Zhang, Y. Jin, M. R. Plummer, S. Pooyan, S. Gunaseelan, P. J. Sinko, *Mol. Pharm.* **2009**, *6*, 836-848; c) D. M. Copolovici, K. Langel, E. Eriste, U. Langel, *ACS Nano* **2014**, *8*, 1972-1994.

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

[19] a) S. Elliott, E. Pham, I. C. Macdougall, *Exp. Hematol.* **2008**, *36*, 1573-1584; b) A. Sanchez-Fructuoso, L. Guirado, J. C. Ruiz, V. Torregrosa, E. Gonzalez, M. L. Suarez, R. Gallego, *Transplant. Proc.* **2010**, *42*, 2931-2934.

[20] 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. Control. Release* **2010**, *145*, 306-313.

[21] P. L. Turecek, M. J. Bossard, F. Schoetens, I. A. Ivens, *J. Pharm. Sci.* **2016**, *105*, 460-475.

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

[23] 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. Nanomedicine* **2011**, *6*, 1217-1227.

[24] 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.

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

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

[27] 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.

[28] 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.

국문초록

생명현상은 복잡하고 정교한 신호전달 체계로 이루어져 있기 때문에, 원하는 특정 생체분자의 정밀한 관찰은 어려운 것으로 생각되어 왔다. 생체 분자를 모니터링하거나 동정해내기 위해서는, 발광체나 특정 작용기와 같은 기능기를 생체분자에 연결해야한다. 고전적으로 생체분자와 리간드는 소수성 상호작용을 통해 연결되어 왔으나, 그 결합은 충분히 강하지 않기 때문에, 생체분자에 대한 공유결합 형성 방법이 개발되어 왔다.

구리(I) 이온은 아자이드와 알카인 사이의 1,3-쌍극부가환화 반응의 촉매로 작용한다는 사실이 알려진 이후, 확장 유전 암호를 이용하여 아자이드 작용기를 단백질 내부에 유전적으로 도입하는 방법이 개발되었다. 이 두 가지 사실을 기반으로, 단백질에 형광체나 항암제 등의 인공기능기를 연결하는 것이 가능해졌다. 특히 생체분자의 형광 표지는 시공간적으로 자유롭게 생체분자를 관찰할 수 있게 한다는 점에서 의미가 있다. 즉, 화학생물학 분야에서 생체분자 표지방법의 개발과 형광 표지분자의 합성 두 가지 모두 중요한 과제라 할 수 있다.

이 박사학위논문은 1) 생체 적용가능한 형광물질을 합성하는 방법론의 개발과 2) 생체분자 내 타이로신의 선택적 표지법 개발을 포함한다.

귀금속 촉매를 이용한 1,3-부가환화 반응과 팔라듐 촉매를 이용한 탄소-수소 활성화 반응을 통해 인돌리진온 기반 형광체를 효율적으로 합성하는 방법론을 제시하였다. 이 형광체의 광물리적 성질은 7, 9번 탄소 치환기에 의해 결정되므로, 다양한 치환기의 도입이 라이브러리 합성 면에서 중요하다. 새로운 합성경로를 통해 9번 탄소 치환기의 도입을 합성 후반부에 위치시켰으며, 기존엔 불가능했던 바이오센서를 합성하였다.

설페이트 클릭 반응을 생체 분자에 적용하여 타이로신 선택적 생체분자 표지법을 개발하였다. 6가 황-불소 교환 반응 (SuFEx)은 방향족 플루오로설페이트와 방향족 실릴에테르 사이의 반응으로 알려졌으나, 방향족 수산화 음이온을 첫 출발물질로 사용한 바는 없었다. 본 연구를 통해 방향족 플로오로설페이트가 단백질 내에 존재하는 다양한 아미노산 중 타

이로신에 대해서만 특이적으로 반응하는 조건을 확보하였다. 나아가, TAT 47-57 펩타이드 내에 존재하는 다양한 아미노산 잔기 중 타이로신에만 선택적으로 형광체를 도입하였으며, erythropoietin 단백질의 외부에 노출된 49번 타이로신에 폴리 에틸렌 글리콜을 표지하였다.

주요어 : 형광, 탄소-수소 활성화, 생체연결, 타이로신, 황-불소 교환반응
학 번 : 2011-20308