



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

이학박사학위논문

Rational Design and Optimization of Novel
Bioprobes and Bioactive Small Molecules
for Exploring Biological Systems

생명 현상의 탐구를 위한 새로운 바이오프로브 및 생리
활성 저분자 화합물의 합리적 디자인 및 최적화

2017년 2월

서울대학교 대학원

화학부 유기화학전공

구 자 영

Abstract

Rational Design and Optimization of Novel Bioprobes and Bioactive Small Molecules for Exploring Biological Systems

Ja Young Koo

Department of Chemistry

College of Natural Science

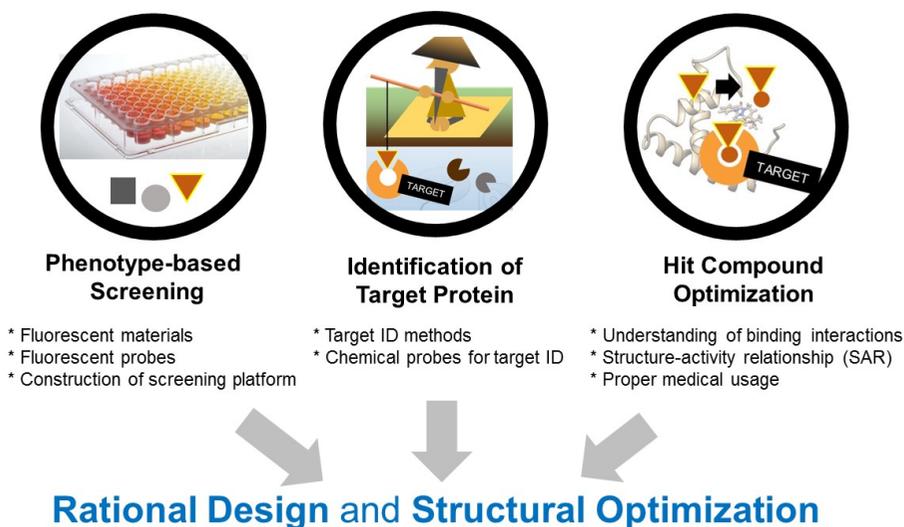
Seoul National University

When we observe microscopic or macroscopic changes in biological systems including human, chemical interactions underlie in every fundamental region. Prior to observing the invisible microscopic world in human body, development of chemical tools for monitoring or perturbing biological systems is necessary. Including many kinds of medicines on the market, small molecules modulators are well-known toolkits for discovering the inside of humans and the living organisms.

Phenotype-based screening has emerged as a promising starting line to discover novel small-molecule drug candidates. In the recent few decades, most

of the newly discovered first-in-class drugs were discovered by phenotype-based approach, which has been the best way to reach the goal.¹ However, this attractive approach has its own limitation. Usually when we found out a potent candidate by the phenotype-based approach, there are more lengthy ways to identify its actual biological mechanism.^{2,3} Target protein identification is a significant hurdle in phenotype-based drug discovery. Phenotype-based discovery of novel bioactive molecules and their target protein identification are facilitated by useful chemical tools. For this purpose, design and optimization of the chemical tools are crucial for expanding our knowledge about unclear and challenging biological events, including incurable diseases or unknown various metabolic pathways.

In general, phenotype-based drug discovery platform consists of (1) phenotype-based discovery of hit compounds, (2) target protein identification of the selected molecules and (3) further hit-to-lead optimization. In this whole drug discovery process, designing proper chemical toolkits for each step has to be preceded. They could be prepared by rational molecular design; Fluorophores and fluorescent probes for monitoring phenotypic changes, probes for identifying the target of a bioactive compound, and development of improved candidates through structure-activity relationship study. Systematic design of biochemical tools is explained in three parts as following.



Scheme 1. Strategy to discover first-in-class drug candidates with phenotype-based screening, target identification process and hit-to-lead optimization.

Part 1 describes the design and construction of naphthalene-based two-photon fluorophores. Fluorescent organic compounds are useful research tools for visualizing biological events including phenotypes in biological systems. Starting from a naphthalene-structure based fluorescent compound, acedan, readily preparable two-photon emission fluorophores were designed with full visible-color coverage. In advance, a two-photon hydrogen peroxide (H_2O_2) sensor was successfully constructed with the turn-on property of red fluorescence.

Part 2 describes the design and optimization of target identification probes for affinity-based 1D or 2D target identification methods. First section of Part 2, FITGE-based target ID and the strategy for target ID probe design was

described with two successful target ID cases. Second section of Part 2, specific binding proteins with each photoaffinity linker in target ID probes were figured out, which proteins could disrupt efficient deconvolution by nonspecific binder of target ID probes. Third section of Part 2, with the change of molecular shapes in target ID probes, length and linearity of target ID linkers could affect the level of their nonspecific protein bindings.

Part 3 describes the structural optimization of hit compounds based on identified biological targets. First section of Part 3, primary computational screening and further biological evaluation was performed to discover a natural product, eryvarin H, with $ERR\gamma$ inverse agonistic function. After total synthesis and derivatization, we finally discovered and synthesized novel $ERR\gamma$ inverse agonists from nature. In the second section of Part 3, treatment for sepsis pathogenesis was successfully done by HMGB1 modulator, ICM. From the structure-activity relationship (SAR) of ICM through HMGB1, compounds **2j** and **2l** with improved anti-inflammatory effects were optimized and successfully ameliorated sepsis pathogenesis in mouse model.

In conclusion, design and optimization of bioprobes and bioactive small molecules in this dissertation suggest us a decisive answer for not only discovering novel first-in-class drug candidates, but also lifting the veil of unknown biological phenomena.

References

- (1) Swinney, D.C.; Anthony, J. *Nat. Rev. Drug Discov.* **2011**, *10*, 507–519.
- (2) Terstappen, G. C.; Schlüpen, C.; Raggiaschi, R.; Gaviraghi, G. *Nat. Rev. Drug Discov.* **2007**, *6*, 891–903.
- (3) Schenone, M.; Dančík, V.; Wagner, B.K.; Clemons, P.A. *Nat. Chem. Biol.* **2013**, *9*, 232–240.

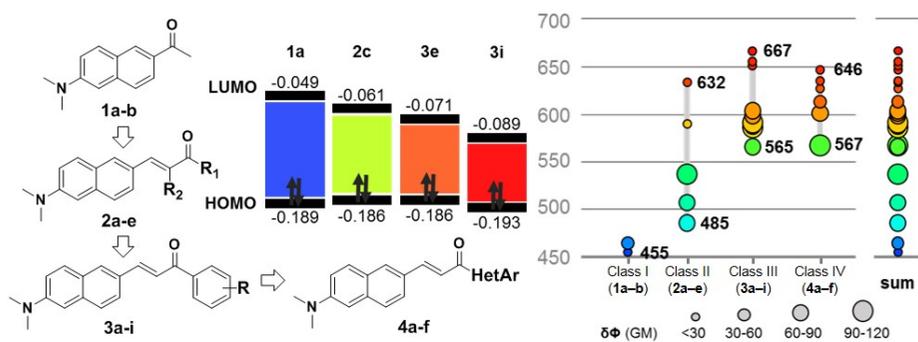
Keywords: Rational Design, Optimization, Small molecule, Hit compound, Drug discovery, Phenotype-based screening, Fluorophore, Fluorescent probe, FITGE, Target identification, Photoaffinity linker, Molecular shape, Eryvarin H, Estrogen-related receptor γ , Inverse agonist, Sepsis, High mobility group box protein, Anti-inflammation

Student Number: 2010-23080

Table of Contents

Abstract	i
Table of Contents	vi
Part 1. Design and Construction of Naphthalene-based Two-photon	
Fluorophores	1
1.1. Readily Accessible and Predictable Naphthalene-based Two-photon Fluorophore with Full Visible-color Coverage	2
Part 2. Design and Optimization of Target Identification Probes	
.....	45
2.1. Probe Design for Target Protein Identification of Bioactive Small Molecules: Design Strategy and Case Studies	46
2.2. Investigation of Specific Binding Proteins to Photoaffinity Linkers for Efficient Deconvolution of Target Protein	87
2.3. Nonspecific Protein Labeling of Photoaffinity Linkers Correlates with Their Molecular Shapes in Living Cells	131
Part 3. Structural Optimization of Hit Compounds based on Identified	
Biological Targets	164
3.1. Total Synthesis of Eryvarin H and its Derivatives as Estrogen- related Receptor γ (ERR γ) Inverse Agonists	165
3.2. Treatment of Sepsis Pathogenesis with High Mobility Group Box Protein 1 (HMGB1)-regulating Anti-Inflammatory Agents	198

Part 1. Design and Construction of Naphthalene-based Two-photon Fluorophores



1.1. Readily Accessible and Predictable Naphthalene-based Two-photon Fluorophore with Full Visible-color Coverage

1.1.1. Abstract

Herein we reported 22 acedan-derived two-photon fluorophores with synthetic feasibility and a full coverage of visible color emission. Their emission wavelengths were well predicted by *in silico* analysis, which enabled us to visualize multicolor images by two-photon excitation with single wavelength, and to design a turn-on two-photon fluorescence sensor for endogenous H₂O₂ in Raw 264.7 macrophage and rat brain hippocampus *ex vivo*.

1.1.2. Introduction

To explore the deep inside of biological systems, fluorescent organic molecules have been essential and powerful materials as a visualization tool due to their ease of handling, low cost, high sensitivity, and bypassing radioactive materials.¹ As bioimaging technologies have been developed, two-photon microscopy has emerged in a last few decades to overcome the limitation of one-photon fluorescence microscopy. By virtue of near-IR light excitation, two-photon organic fluorophores have their own advantages for bioimaging such as the increased penetration depth for *ex vivo* imaging,

tolerance toward photo-damages and photo-bleaching, and localized excitation.²

Among two-photon organic fluorophores, acedan (**1a**)—naphthalene-based fluorophore containing dipolar donor-bridge-acceptor system(D- π -A)^{3,4}—has been the most popular structural motif for the biological applications and has been utilized in monitoring various biological events such as changes in metal ions,⁵ pH,⁶ reactive oxygen species,⁷ and other metabolites.⁸ Due to the fact that acedan emits bright fluorescence by single-photon excitation as well as two-photon excitation despite its relatively simple structure, its analogues have been developed for two-photon bioimaging.

However, structure-photophysical property relationship (SPPR) studies of acedan-derived fluorophores have not yet been pursued systematically. This SPPR information is important for designing bioprobes to visualize biological systems with multicolor windows and helpful to select appropriate two-photon fluorophores with desired photophysical properties for certain biological applications. Along with expanding our knowledge of two-photon fluorophores derived from naphthalene-based D- π -A system, we focused on a rational approach for designing bright two-photon fluorophores with synthetic feasibility as well as a full visible-color coverage of emission wavelength (λ_{em}).

1.1.3. Results and Discussion

Looking in-depth, acedan **1a** is composed of simple naphthalene core skeleton with two substituents, 2-acetyl and 6-dimethylamino groups with opposite electronic characters, at the opposite directions. The photophysical

property of this common structural motif can be explained by intramolecular charge transfer (ICT)⁹ process; electronic perturbation through p-electron linkage from electron-donating part (6-dimethylamino group) to electron-withdrawing part (2-acetyl group) is observed at the calculated excited states of **1a**, which makes it a typical example of organic fluorophores containing the dipolar D- π -A system (Figure 1a). As **1a** emits fluorescence at blue-to-green colored region (455 nm in DMF and 515 nm in water⁴), we aimed its structural modification for bathochromic shift of emission wavelength. Before the actual synthesis of fluorophores derived from **1a**, we carried out *in silico* analysis to obtain the structural insight for the rational modification of **1a** to minimize trials and errors. In fact, our previous studies clearly demonstrated that the emission wavelength of organic fluorophores are inversely correlated to the energy gap between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO).^{10,11} Their molar absorptivity (ϵ) is also positively correlated with S₀-S₁ oscillator strength.^{11,12} Based on this guideline, we designed a series of analogues of **1a** with a decrease of HOMO-LUMO energy gap to achieve red-shifted emission wavelengths.

When we introduced various substituents, instead of dimethylamino group, at the C-6 position of **1a**, we observed wider HOMO-LUMO energy gaps (Figure S1.1.1). In contrast, the introduction of additional π -conjugation system, various functional groups, and substituted aryl/heteroaryl groups at the C-2 position of **1a** led to narrower HOMO-LUMO energy gaps. As shown in Figure 1.1.1a, a simple insertion of olefin unit between naphthyl and carbonyl group resulted in **2c** with 82-nm red shift in the predicted emission wavelength. The

further extension of π -conjugation system via additional phenyl ring beside carbonyl group in **2c** generated **3e**, which was predicted to have an additional 73-nm red shift. 4-cyanophenyl-containing **3i** predicted to show additional 85-nm bathochromic shift from **3e**. Interestingly, **3i** was predicted to have 212 nm red-shifted emission wavelength compared to that of **1a**. The increase of electron-withdrawing natures at the C-2 position of **1a** via an extension of π -conjugation system or an introducing more electron-withdrawing substituents lowers the LUMO energy (Figure 1.1.1b). Due to the fact that the lowering of LUMO energy strongly influences the ICT process of dipolar D- π -A system, these rational modifications of **1a** caused larger shifts of orbital lobes between HOMO and LUMO (Figure 1.1.1c).

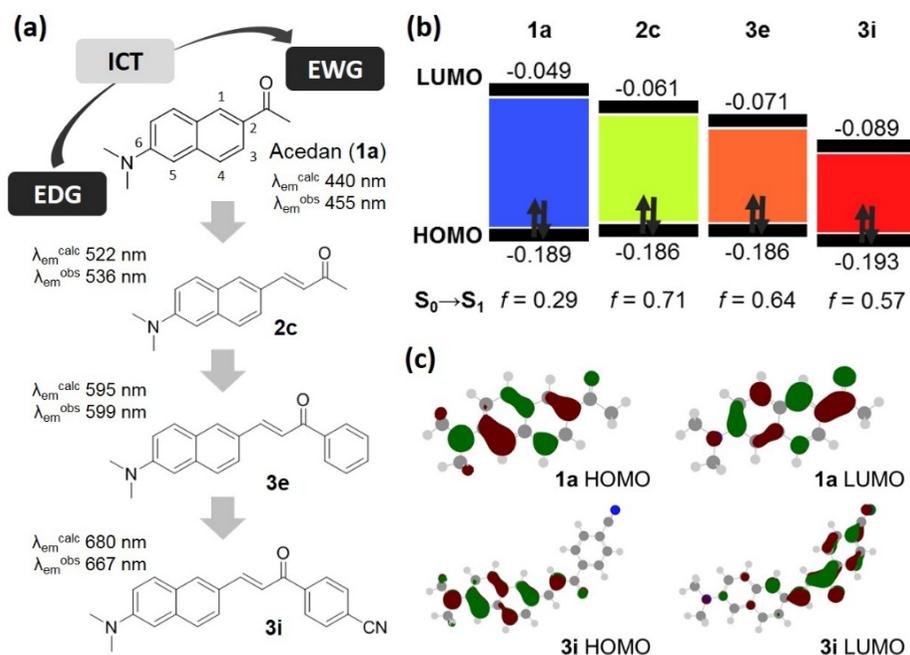
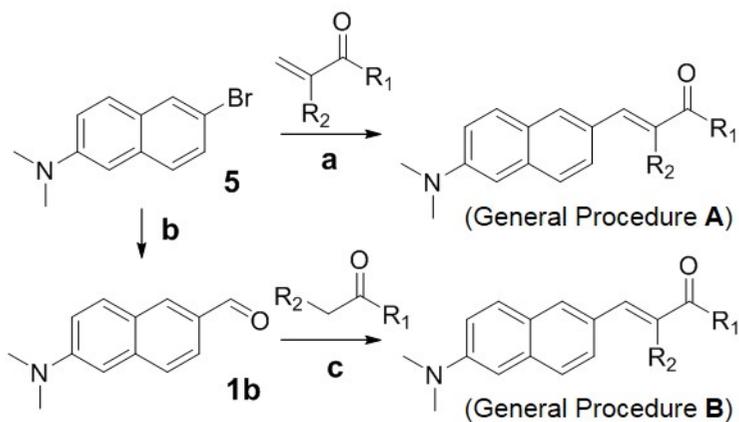


Figure 1.1.1. Design principle of naphthalene-based fluorophores from **1a** to **2c**, **3e**, and **3i**. a)

Designing streamline of naphthalene-based fluorophores from blue to red fluorescence. λ_{em}^{calc} : Predicted emission wavelength (nm) (see Table S1); λ_{em}^{obs} : Observed emission wavelength (nm). b) HOMO and LUMO energy levels and energy gaps of representative fluorophores (**1a**, **2c**, **3e**, and **3i**). c) Molecular orbital lobes of **1a** and **3i** at HOMO and LUMO states. Larger shift of orbital lobes was observed in **3i** than **1a**.

For the practical usage of organic fluorophores, it is essential to have a readily accessible synthetic route. In our naphthalene-based two-photon fluorophores, all desired compounds were prepared within one- or two-step synthesis from previous reported intermediate **5**. As shown in Scheme 1, Pd-mediated Heck-type coupling reaction of **5** with various α,β -unsaturated carbonyl compounds yielded the desired fluorophores in a single step (General Procedure A). The identical set of organic fluorophores can be achieved using aldehyde-containing **1b** obtained by lithiation of **5** with *n*-BuLi and the subsequent formylation via the substitution with DMF.¹³ The resulting intermediate **1b** was transformed to the desired fluorophores via simple aldol condensation (General Procedure B).

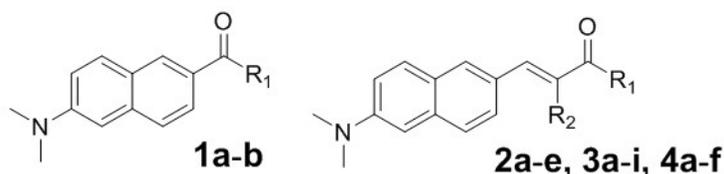


Scheme 1.1.1. General synthetic schemes for designed naphthalene-based fluorophores; Desired molecules were synthesized through general procedure **A** (**2a** and **2b**) or **B** (**2c–e**, **3a–i**, and **4a–f**) depending on their substituents. Reaction conditions: a) NaHCO₃, Pd(PPh₃)₂Cl₂, DMF, 140 °C, 8 h; b) n-BuLi, DMF, THF, –78 °C → 0 °C, 3 h; c) NaOH, ethanol, water, r.t., 24 h or acetic acid, piperidine, ethanol, reflux, 8 h.

Using this synthetic strategy, we prepared four classes of naphthalene-based two-photon fluorophores. Class I consists of original acedan (**1a**) and formyl-substituted analogue **1b**. Two-photon organic fluorophores in Class II, III, and IV were prepared by either General Procedure A or B. Class II contains olefin-inserted acedan (**2c**) and their analogues with either electron-donating groups (**2a**, **2b**) or withdrawing groups (**2d**, **2e**). As shown in Table 1.1.1, we observed the longest emission wavelength in the case of electron-deficient trifluoromethyl-containing **2e** with additional 100-nm red shift from **2c**. Electron-withdrawing cyano moiety at the R₂ position also caused 80-nm red shift from **2b** to **2d**. To further extend the π -conjugation system, we synthesized **3a–i** as Class III by introducing various aryl moieties at the R₁ position. As expected, the phenyl-containing **3e** has 50-nm longer emission wavelength than methyl-containing **2c**. It is worth mentioning that the electron poorness in aryl moieties at the R₁ position can make more efficient push-pull interaction with the dimethylamino group, which led to gradual red shifts of emission. Fluorescent emission wavelength of these fluorophores effectively covers full visible-color range and their brightness is comparable or better than original acedan **1a** with two-photon ($\delta\Phi$) excitation (Table 1.1.1 and size of circle in Figure 1.1.2a) as well as one-photon ($\epsilon\Phi$) excitation (Figure S1.1.3). Although

it was predicted, it is disappointing to observe that the fluorescence quantum yield was decreased upon the increase of emission wavelength in our naphthalene-based fluorophores with λ_{em} greater than 600 nm. The observed emission wavelength in each class of fluorophores was well-matched with the calculated energy gap between HOMO and LUMO (Table S1.1.1). In fact, electronic character of acetyl group in **1a** was significantly altered by the elongation of π -conjugated system via an insertion of olefin unit or the further attachment of aromatic ring, which lowers energy levels of LUMO in Class II (**2a–e**), III (**3a–i**), and IV (**4a–f**), compared to **1a** (Figure 1.1.2b). The systematic alteration of electronic deficiency in aromatic rings in Class III, quantified by Hammett constant (σ_p), showed the linear correlation with the inverse of fluorescence emission wavelength (Figure 1.1.2c and S1.1.4). In addition, the observed emission wavelength of each fluorophore showed an excellent linear correlation with the inverse of HOMO–LUMO energy gap (Figure 1.1.2d), which confirmed the predictability of emission wavelength in our two-photon fluorophores. Moreover, there is a possibility for two-photon multicolor fluorescence imaging with our fluorophores in a visible light range when we apply appropriate light source. With double excitation with the IR-range laser source at 800 nm, we simultaneously observed blue-to-red two-photon fluorescence (Figure 1.1.2e). It is worth mentioning that our representative fluorophores **1a**, **2a**, **2b**, **2c**, **4a**, **3e**, and **4e** showed their unique two-photon cross sections upon double excitation at 760 and 800 nm, respectively (Figure 1.1.2f). This unique feature of a full visible-color coverage of two-photon fluorescence with double excitation at a single fixed wavelength

can be an essential element for the further application into multiplexed full-color bioimaging.



Cpd.	R ₁	R ₂	$\lambda_{\text{ex}}^{\text{a}}$	λ_{em}	$\delta\Phi$	$\lambda_{\text{ex}}^{\text{b}}$	ϵ	Φ	$\epsilon\Phi$
1a	CH ₃		740	455	22	354	47k	0.62	29k
1b	H		750	463	30	373	31k	0.84	26k
2a	N(CH ₃) ₂	H	740	485	79	361	25k	1.00	25k
2b	OCH ₃	H	760	507	74	382	52k	0.96	50k
2c	CH ₃	H	750	536	92	388	60k	0.93	56k
2d	OCH ₃	CN	840	591	7.5	455	16k	0.15	2.4k
2e	CF ₃	H	900	632	1.2	461	33k	0.04	1.3k
3a	4-NH ₂ Ph	H	820	565	74 ^c	414	63k	0.40	25k
3b	3,4-(CH ₃ O) ₂ Ph	H	840	584	111	422	50k	0.57	29k
3c	4-CH ₃ OPh	H	840	586	88	423	37k	0.51	19k
3d	4-CH ₃ Ph	H	840	590	100	420	29k	0.49	14k
3e	Ph	H	840	599	71	423	37k	0.46	17k
3f	4-FPh	H	840	602	74	423	35k	0.46	16k
3g	4-CF ₃ Ph	H	900	650	1.0	436	17k	0.05	0.9k
3h	4-CH ₃ COPh	H	940	657	0.3	444	33k	0.006	0.2k
3i	4-CNPh	H	940	667	0.3	449	25k	0.012	0.3k
4a	pyrrol-2-yl	H	820	567	98	412	24k	0.59	16k
4b	furan-2-yl	H	840	601	89	427	20k	0.61	12k
4c	thiophen-2-yl	H	860	612	52	431	40k	0.38	17k
4d	pyrid-3-yl	H	840	625	18	435	36k	0.18	7.9k
4e	pyrid-2-yl	H	840	633	6.1	437	34k	0.08	2.4k
4f	pyrid-4-yl	H	840	646	0.9	441	52k	0.03	0.5k

Table 1.1.1. Two-photon and one-photon photophysical properties of **1a–b**, **2a–e**, **3a–i**, **4a–f** in DMF; λ_{ex} = fluorescence excitation wavelength (nm); λ_{em} = emission wavelength (nm); $\delta\Phi$ = two-photon cross section (GM, $10^{-50} \text{ cm}^4 \cdot \text{s} \cdot \text{photon}^{-1}$); ϵ = molar extinction coefficient ($\text{L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$); Φ = absolute quantum yield. ^aTwo-photon excitation; ^bOne-photon excitation; ^cLocal

maximum point of two-photon emission fluorescence. The global maximum cross section of **3a** is 129 GM by two-photon excitation at 750 nm.

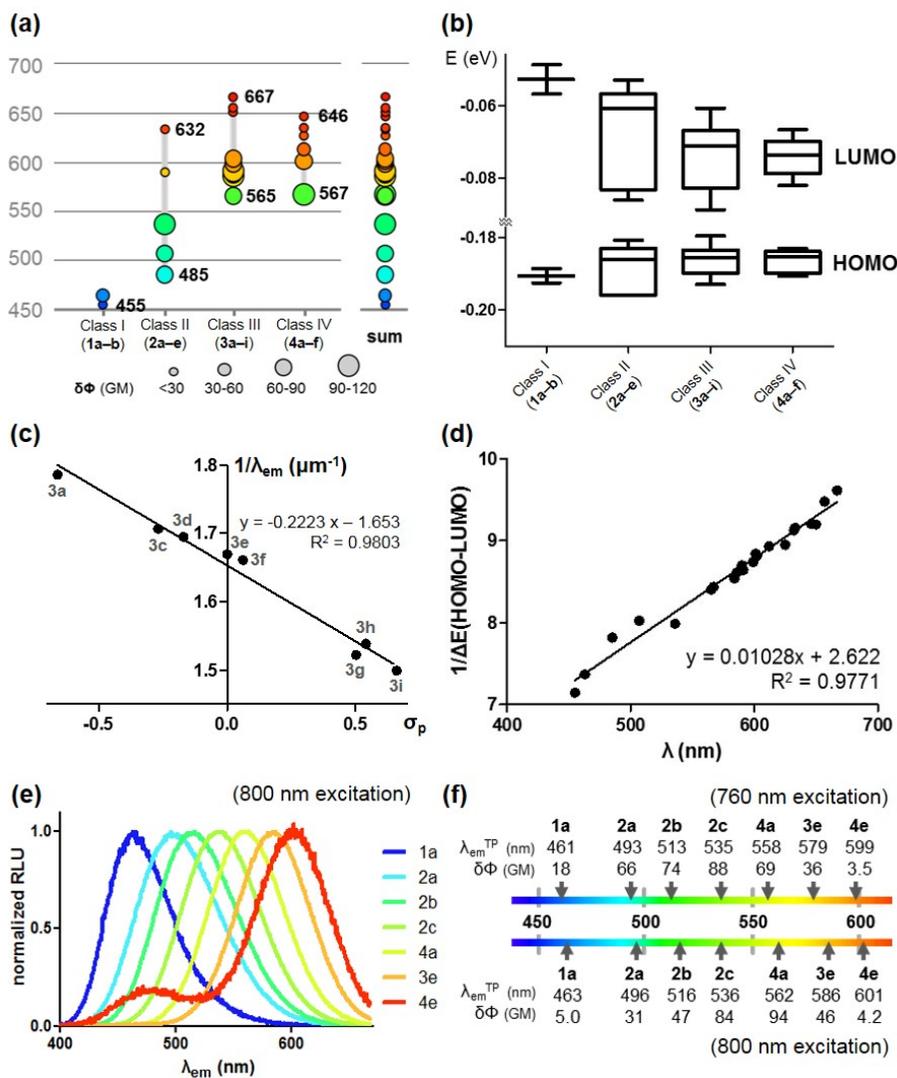


Figure 1.1.2. Correlation between predicted and observed photophysical property data of all 22 prepared fluorescent compounds. a) Two-photon cross section ($\delta\Phi$) and emission wavelength of **1a–b**, **2a–e**, **3a–i**, **4a–f**; size of circle represents two-photon cross section; y-value and color of circle represent the emission wavelength. b) HOMO and LUMO energy levels of **1a–b**, **2a–e**, **3a–i**, **4a–f** were plotted in vertical box plot. c) Tendency between Hammett constant (σ_p) and

emission wavelength of **3a–i** as the substituent changes. d) Linear regression between emission wavelength (nm) and inverse of HOMO–LUMO energy gap (1/eV). e) Normalized fluorescence intensity of **1a**, **2a**, **2b**, **2c**, **4a**, **3e**, and **4e** originated from double excitation with 800 nm. f) Emission wavelength (nm) and cross sections ($\delta\Phi$) of two-photon fluorescence of **1a**, **2a**, **2b**, **2c**, **4a**, **3e**, **4e**, and **3i** excited by 760 or 800 nm, respectively.

For the application of our emission-tunable two-photon fluorophores, we designed and synthesized a fluorogenic H₂O₂ sensor **3j** by one-step reaction using intermediate **5**. As shown in Figure 1.1.3a, **3j** and H₂O₂-reacted product **3k** are categorized in Class III because they contain substituted phenyl moieties at the R₁ position. Furthermore, their properties are suitable for the rational design of red-shifted two-photon sensors. Based on the calculated HOMO–LUMO energy gap of **3j** and **3k**, the predicted emission wavelengths using linearity in Figure 2d were within a 20-nm range of deviation from the observed emission wavelength, 598 nm and 575 nm, respectively. When we measured the brightness of one-photon ($\epsilon\Phi$: 8100, 15000 L·mol⁻¹·cm⁻¹) and two-photon ($\delta\Phi$: 13, 71 GM) excited fluorescence of **3j** and **3k** at their emission maxima (Figure S5), respectively, we envisioned that the functional group transformation by oxidative stress from borono (**3j**) to hydroxyl group (**3k**) in our naphthalene-based fluorophore can make novel fluorogenic sensor for hydrogen peroxide, an endogenous reactive oxygen species (Figure 1.1.3a). To test our hypothesis, **3j** was treated with H₂O₂ in test tubes, which converts **3j** to **3k** in a dose- and time-dependent manner (Figure S1.1.6). As shown in Figure 1.1.3b, we observed 11-fold increase of fluorescence intensity at 641 nm upon treatment with hydrogen peroxide. Then, we examined whether fluorogenic

H₂O₂ sensor **3j** is applicable in living cells. After pre-incubation with **3j**, Raw 264.7 macrophage cells were treated with H₂O₂ or vehicle, which led to the increase of fluorescence intensity in cytoplasm (Figure 1.1.3c). Once we confirmed the fluorogenic characteristics of **3j** toward exogenous H₂O₂ in the cellular system, we further tested whether the endogenous H₂O₂ could be detected by **3j**. Raw 264.7 cells were incubated with **3j** followed by the treatment with lipopolysaccharide (LPS), which stimulates the immune response in macrophage cells and triggers the release of various toxic chemokines, including hydrogen peroxide and nitric oxide. As shown in Figure 1.1.3d, the LPS-treated cells showed a gradual increase of fluorescence intensity in cytosol while vehicle-treated cells did not.

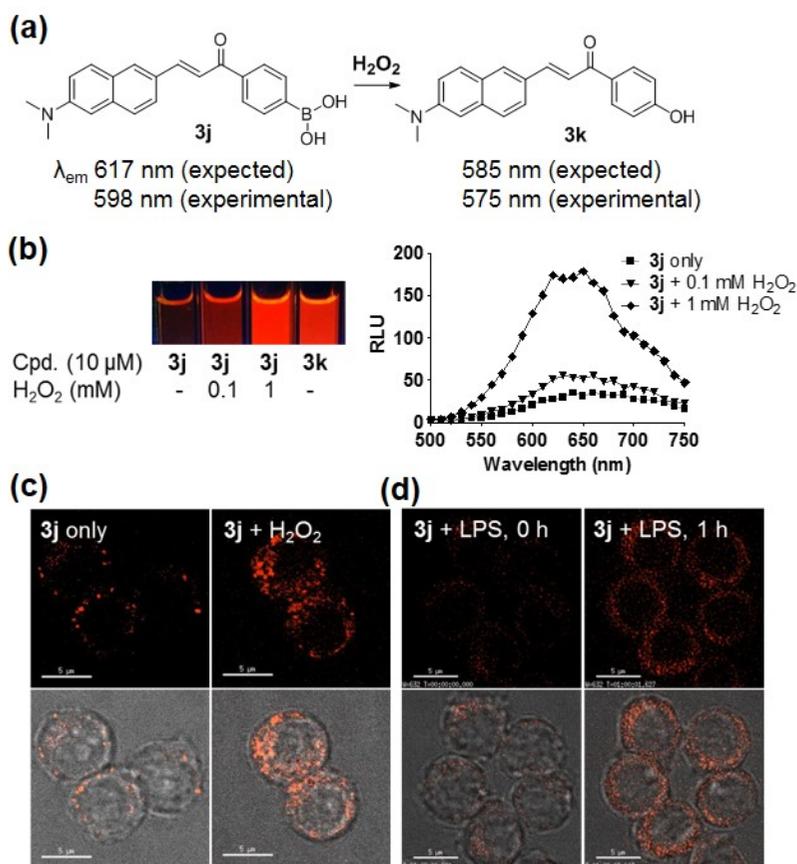


Figure 1.1.3. Design strategy and biological application of H₂O₂ fluorescent sensor. a) H₂O₂ sensor **3j** was prepared in a single step from **5**. Emission wavelengths of **3j** and **3k** were well matched with observed data within a 20 nm deviation range; b) **3j** was transformed to **3k** upon *in vitro* treatment of exogenous H₂O₂. (11-fold increase at 641 nm, full conversion was checked by LC-MS). c) Fluorescence images of **3j** in Raw 264.7 cells upon treatment of exogenous H₂O₂. After 20-min treatment with 2 μ M **3j**, Raw 264.7 cells were incubated in the presence (right) or absence (left) of 200 μ M H₂O₂ for 1 h. d) Raw 264.7 cells were pretreated with 1 μ M **3j** for 20 min, and then time-lapse images were taken upon addition of LPS (500 ng/mL). Fluorescence images (red) were taken from λ_{ex} 438 nm/ λ_{em} 632nm. Scale bar is 5 μ m.

Finally, we investigated the utility of **3j** for tissue imaging applications by using two-photon microscopy (TPM). Two-photon excited fluorescence facilitates to visualize organs *ex vivo* due to the lengthened penetration depth. Therefore, we applied this H₂O₂ sensor into the two-photon *ex vivo* imaging. A fresh rat hippocampal slice was incubated with **3j** and the TPM images were collected upon excitation with 820 nm with femtosecond pulses. The images displayed weak two-photon excited fluorescence (TPEF) intensities in the CA1 and CA3 regions (Figure 1.1.4a and 1.1.4d). The TPEF intensities in both regions increased after treatment with H₂O₂ (Figure 1.1.4c, 1.1.4f, and 1.1.4h). Similar result was observed when the tissue was stimulated with phorbol-12-myristate-13-acetate (PMA), which induces cellular H₂O₂ generation through inflammation processes (Figure 1.1.4b and 1.1.4e). Hence, our probe reports the increase of endogenous H₂O₂ in living tissue. The successful application of hydrogen peroxide sensor **3j** in the cellular imaging by one-photon excitation as well as *ex vivo* imaging by two-photon excitation clearly confirmed that our naphthalene-based fluorophore can be utilized for the rational development of simple and predictable fluorescence sensors.

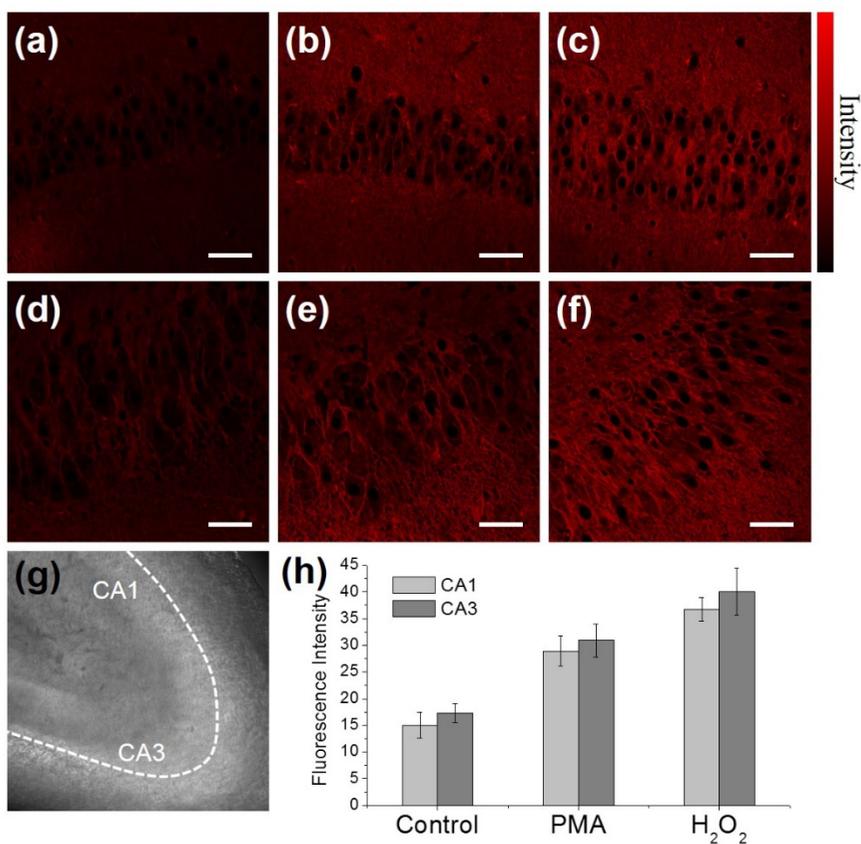


Figure 1.1.4. TPM images of a rat hippocampal slice stained with (a, d) 5 μM **3j**. A rat hippocampal slices were pretreated with (b, e) PMA (5 $\mu\text{g mL}^{-1}$) for 30 min and (c, f) 1 mM H₂O₂ for 30 min before labeling with 5 μM **3j**. (a–c) CA1 and (d–f) CA3 regions. (g) Bright-field image of the CA1 and CA3 regions at 10x magnification. (h) Average TPEF intensities in (a–f). The TPEF were collected at 550–700 nm upon excitation at 820 nm with fs pulse. Scale bars: (a–f) 48 μm .

1.1.4. Conclusion

In summary, we systematically constructed 22 naphthalene-based two-photon fluorophores with a full coverage of visible light emission. Based on ICT-based structural motif in **1a**, the emission wavelength of our fluorophores can be predicted with high accuracy due to its linear correlation with the inverse of HOMO–LUMO energy gap. The structure–photophysical property relationship study of our two-photon fluorophore guided the rational development of fluorogenic H₂O₂ sensor **3j**, which exhibits the turn-on fluorescence upon exposure to LPS-mediated endogenous H₂O₂ as well as exogenous H₂O₂ in Raw 264.7 macrophage. Finally, **3j** was effectively applied to two-photon *ex vivo* imaging in a rat hippocampal slice, which was resulted in the visualization of endogenous H₂O₂ with red-colored two-photon fluorescence signal.

1.1.5. Experimental Section

Computational Calculation

Prior to the optimization of compounds' structure, the most stable conformer of each compound was calculated by Vconf2.0 using Tork conformational analysis method.¹⁴ Energy minimization, structural optimization, and HOMO/LUMO energy level calculation of each compound were performed by DMol3 in Materials Studio[®] 4.2 program [Accelrys Software, Inc.]. BLYP density functional, exchange-correlation function of PBE (Perdew, Burke, and

Ernzerhof) was used with DNP (double numerical basis set with polarization). S_0 – S_1 energy gap and oscillator strength of the compounds were calculated by Gaussian W09. Time-dependent density functional theory (TD-DFT) calculation was performed with B3LYP density functional and 6-31G(d) basis set.

Measurement of Molar Extinction Coefficients (ϵ) and Absolute Quantum Yields (Φ)

Absorption spectra and molar extinction coefficients (ϵ) at the absorption maxima of our fluorophores were measured by UV-VIS spectrophotometer UV-1650PC [Shimadzu, Japan]. Emission wavelengths and spectra were measured by Cary Eclipse Fluorescence spectrophotometer [Varian Associates]. Absolute quantum yields (Φ) of compounds was measured by QE-2000 [Otsuka Electronics].

Measurement of Two-photon Cross Sections ($\delta\Phi$)

The two-photon cross section (δ) of fluorophores were determined by using a femtosecond (fs) fluorescence measurement technique as previously reported.¹⁵ Each two-photon fluorophore was dissolved in DMF, and measured at the concentration of 5.0×10^{-6} M. Two-photon excited fluorescence intensity was measured from 720 nm to 940 nm by using Rhodamine 6G in methanol as the reference, whose two-photon properties were well characterized in the previous literature.¹⁶

The intensities of the two-photon excited fluorescence spectra of the reference

compound and our sample emitted at the same excitation wavelength were determined. Each of the two-photon absorption (TPA) cross section was calculated by using $\delta = \delta_r(S_s\Phi_r\varphi_r c_r)/(S_r\Phi_s\varphi_s c_s)$, where the subscripts *s* and *r* stand for the sample and reference molecules. The intensity of the signal collected by a CCD detector was denoted as *S*. Φ is the fluorescence quantum yield. φ is the overall fluorescence collection efficiency of the experimental apparatus. The number density of the molecules in solution was denoted as *c*. δ_r is the TPA cross section of the reference molecule.

General Information for Biological Experiments

One-photon fluorescence bioimaging studies were performed using DeltaVision Elite imaging system [GE Healthcare] equipped with a sCMOS camera. Objective lenses are supported by Olympus IX-71 [Olympus] inverted microscope equipped with Plan APO 60X/Oil (PLAPON60×O), 1.42 NA, WD 0.15 mm. DeltaVision Elite uses a solid state illumination system, InSightSSI fluorescence illumination module. Four-color standard filter set [GE Healthcare, 52-852113-003] was used to detect fluorescence signals. For live cell continuous monitoring, CO₂ supporting chamber along with an objective air heater was set and ultimatefocus hardware autofocus module was incorporated to maintain the sample z-position during time-lapse imaging.

H₂O₂ was purchased Samchun chemical (Korea), and Lipopolysaccharides from *Escherichia Coli* 055:B5 purified by phenol extraction (LPS) was purchased from Sigma-Aldrich. Raw264.7 cell lines were obtained from Korean Cell Line Bank. Raw 264.7 cells were cultured in DMEM [GIBCO,

Invitrogen] supplemented with heat-inactivated 10 % (v/v) fetal bovine serum [GIBCO, Invitrogen] 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].

One-Photon Fluorescence Microscopy

Raw 264.7 cells were seeded on cover glass-bottom dish and incubated at 5 % CO₂ incubator at 37 °C overnight. For the exogenous H₂O₂ detection experiment, Raw264.7 macrophage cells were incubated with 2 μM **3j** for 30 min. After replacing media, cells were treated with 200 μM H₂O₂ or vehicle for 1 h. Media was replaced with DMEM not containing phenol red, and then fluorescence images were taken under the microscopy. For the endogenous H₂O₂ detection, Raw264.7 cells were incubated with 1 μM **3j** for 30 min. After replacing media not containing phenol red, LPS (500 ng/mL) or vehicle were added to cells, and fluorescence intensity changes were continuously monitored by Delta Vision for 1 h.

Two-Photon Fluorescence Microscopy (*ex vivo* imaging)

Two-photon fluorescence microscopy images of probe-labeled tissues were obtained with spectral confocal and multiphoton microscopes (Leica TCS SP8 MP) with 10× dry and 40× oil objectives, numerical aperture (NA) = 0.30 and 1.30, respectively. The two-photon fluorescence microscopy images were obtained with a DMI6000B Microscope (Leica) by exciting the probes with a mode-locked titanium-sapphire laser source (Mai Tai HP; Spectra Physics, 80

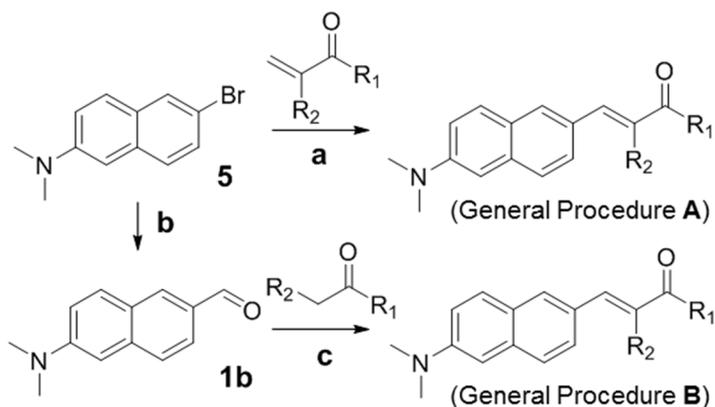
MHz pulse frequency, 100 fs pulse width) set at wavelength 820 nm and output power 2790 mW, which corresponded to approximately 2.70×10^8 mW/cm² average power in the focal plane. Imaging was performed using the live cell incubator systems (Chamlide IC; Live Cell Instrument) for stable cell environment by maintaining proper temperature, humidity and pH over the long term. To obtain images at 550–700 nm range, internal PMTs were used to collect the signals in an 8 bit unsigned 512×512 and 1024×1024 pixels at 400 Hz scan speed, respectively.

Preparation and Staining of Fresh Rat Hippocampal Slices

Rat hippocampal slices were prepared from the hippocampi of 2-weeks-old rat (SD) according to an approved institutional review board protocol. Coronal slices were cut into 400 μm-thicks using a vibrating-blade microtome in artificial cerebrospinal fluid (ACSF; 138.6 mM NaCl, 3.5 mM KCl, 21 mM NaHCO₃, 0.6 mM NaH₂PO₄, 9.9 mM D-glucose, 1 mM CaCl₂, and 3 mM MgCl₂). Slices were incubated with 5 μM probe in ACSF bubbled with 95 % O₂ and 5 % CO₂ for 1 h at 37 °C. Slices were then washed three times with ACSF and transferred to glass-bottomed dishes [NEST] and observed in a spectral confocal multiphoton microscope.

Preparation of compound 2a–e, 3a–k, 4a–f

Preparation and characterization of all new compounds (**2a–e**, **3a–k**, **4a–f**) were described. Preparation and characterization of **1a**,^{17,18} **1b**,¹⁸ and **5**^{18,19} were previously reported.



Reaction conditions: a) NaHCO_3 , $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$, DMF, $140\text{ }^\circ\text{C}$, 8 h; b) $n\text{-BuLi}$, DMF, THF, $-78\text{ }^\circ\text{C} \rightarrow 0\text{ }^\circ\text{C}$, 3 h; c) NaOH , ethanol, water, r.t., 24 h or acetic acid, piperidine, ethanol, reflux, 8 h.

General Procedure A

(a) To a solution of **5** (50 mg, 1 equiv.) and NaHCO_3 (3 equiv.) in anhydrous DMF (5 mL), an appropriate α,β -unsaturated carbonyl substrate (1.5 equiv.) and $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (5 mol %) were added sequentially. The reaction mixture was stirred at $140\text{ }^\circ\text{C}$ for 8 h. The resulting mixture was concentrated, and diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO_4 and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired fluorescent product.

Compounds **2a** and **2b** were prepared by general procedure A.

General Procedure B

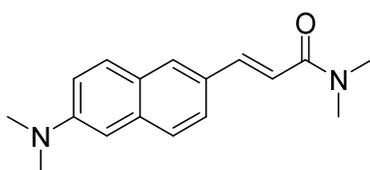
(b) Compound **1b** was prepared by previously reported procedure.¹⁸

(c) *Aldol condensation reaction.* Compound **1b** (50 mg, 1 equiv.) and appropriate carbonyl substrate (1.5 equiv.) was dissolved in ethanol (3 mL). 1:1 mixture of 2N aqueous NaOH solution and ethanol (6 mL) was added to the reaction mixture. The reaction mixture was stirred at room temperature for 24 h, until the starting material was consumed. The resulting mixture was concentrated, and diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired fluorescent product. (*Aldol-1* procedure)

There is an alternative route for this reaction. Compound **1b** (50 mg, 1 equiv.) and appropriate carbonyl substrate (1.5 equiv.) were dissolved in ethanol (3 mL). Acetic acid (1.10 equiv.) and piperidine (1.05 equiv.) were added to the reaction mixture. The reaction mixture was stirred under the reflux condition for 8 h until the starting material was consumed. The resulting mixture was concentrated and diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired fluorescent product. (*Aldol-2* procedure)

Compound **2c–e**, **3a–i** and **4a–f** were prepared by general procedure **B**. Few variations for the reaction completion were described in each compound section.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-*N,N*-dimethylacrylamide (**2a**)



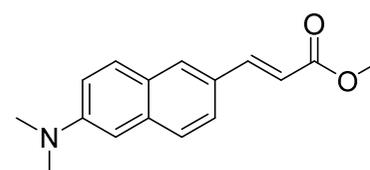
2a was prepared by general procedure **A**.

(Yield: 67%)

$R_f = 0.10$ (1:1 = EtOAc:n-hexane, v/v); ^1H

NMR (400 MHz, CDCl_3) δ 7.77–7.81 (m, 2H), 7.70 (d, $J = 8.8$ Hz, 1H), 7.57–7.62 (m, 2H), 7.14 (dd, $J = 8.8$ Hz, 2.4 Hz, 1H), 6.86–6.93 (m, 2H), 3.19 (s, 3H), 3.08 (s, 3H), 3.07 (s, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 167.3, 149.4, 143.2, 135.9, 129.6, 129.5, 129.1, 126.8, 126.4, 124.2, 116.5, 115.1, 106.1, 40.8, 37.6, 36.1; LRMS(ESI^+) m/z calcd for $\text{C}_{17}\text{H}_{21}\text{N}_2\text{O}$ [$\text{M}+\text{H}$] $^+$ 269.17; Found 269.2.

■ Methyl (*E*)-3-(6-(dimethylamino)naphthalen-2-yl)acrylate (**2b**)



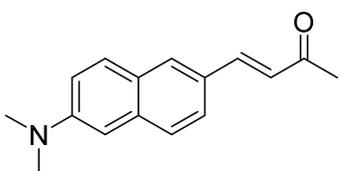
2b was prepared by general procedure **A**.

(Yield: 57%)

$R_f = 0.71$ (1:1 = EtOAc:n-hexane, v/v); ^1H

NMR (400 MHz, CDCl_3) δ 7.81 (d, $J = 16.4$ Hz, 1H), 7.77 (s, 1H), 7.70 (d, $J = 8.8$ Hz, 1H), 7.53–7.63 (m, 2H), 7.14 (dd, $J = 8.8$ Hz, 2.0 Hz, 1H), 6.87 (s, 1H), 6.45 (d, $J = 16.0$ Hz, 1H), 3.82 (s, 3H), 3.08 (s, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 168.1, 149.6, 145.7, 136.3, 130.2, 129.7, 128.0, 126.9, 126.1, 124.1, 116.5, 115.3, 106.0, 51.7, 40.6; LRMS(ESI^+) m/z calcd for $\text{C}_{16}\text{H}_{18}\text{NO}_2$ [$\text{M}+\text{H}$] $^+$ 256.13; Found 256.1.

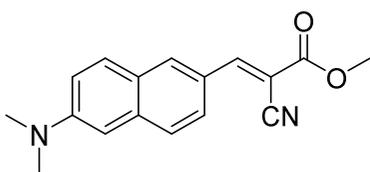
■ (*E*)-4-(6-(Dimethylamino)naphthalen-2-yl)but-3-en-2-one (**2c**)



2c was prepared by general procedure **B**, *Aldol-1* procedure. 2 equiv. of reacting substrate (acetone) was used. (Yield: 43%)

R_f = 0.61 (1:1 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.80 (s, 1H), 7.71 (d, J = 9.2 Hz, 1H), 7.54–7.67 (m, 3H), 7.15 (dd, J = 8.8 Hz, 2.4 Hz, 1H), 6.87 (d, J = 2.0 Hz, 1H), 6.75 (d, J = 16.0 Hz, 1H), 3.09 (s, 6H), 2.40 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 198.6, 149.7, 144.6, 136.5, 130.7, 129.8, 128.0, 127.0, 126.2, 125.2, 124.2, 116.5, 106.0, 40.7, 27.6; LRMS(ESI⁺) m/z calcd for $\text{C}_{16}\text{H}_{17}\text{NO}$ $[\text{M}+\text{H}]^+$ 240.14; Found 240.3.

■ Methyl (*E*)-2-cyano-3-(6-(dimethylamino)naphthalen-2-yl)acrylate (**2d**)

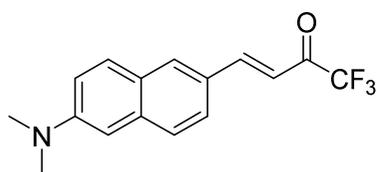


2d was prepared by modified general procedure **B**, *Aldol-1* procedure. SMs were reacted with piperidine (0.1 equiv.) in THF

(5 mL) solution at 70 °C for 12 h. (Yield: 55%).

R_f = 0.64 (1:1 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.28 (s, 1H), 8.20 (s, 1H), 8.09 (dd, J = 8.8 Hz, 1.6 Hz, 1H), 7.76 (d, J = 9.6 Hz, 1H), 7.63 (d, 8.8 Hz, 1H), 7.14 (dd, J = 8.8 Hz, 2.8 Hz, 1H), 6.84 (d, J = 2.4 Hz, 1H), 3.93 (s, 3H), 3.13 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 164.3, 155.7, 150.9, 138.0, 135.3, 131.1, 127.0, 126.3, 125.4, 125.0, 116.9, 116.4, 105.4, 97.8, 53.2, 40.5; LRMS(ESI⁺) m/z calcd for $\text{C}_{17}\text{H}_{17}\text{N}_2\text{O}_2$ $[\text{M}+\text{H}]^+$ 281.13; Found 281.2.

■ (E)-4-(6-(Dimethylamino)naphthalen-2-yl)-1,1,1-trifluorobut-3-en-2-one
(2e)

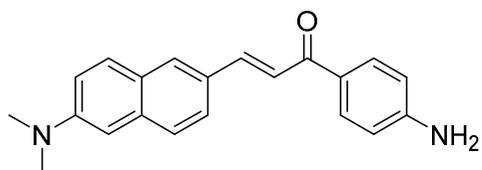


was used. (Yield: 35%)

2e was prepared by modified general procedure **B**, *Aldol-2* procedure. 5 equiv. of reacting substrate, 1,1,1-trifluoroacetone,

$R_f = 0.46$ (1:5 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.08 (d, $J = 15.6$ Hz, 1H), 7.90 (s, 1H), 7.74 (d, $J = 9.6$ Hz, 1H), 7.62 (d, $J = 1.2$ Hz, 1H), 7.15 (dd, $J = 8.8$ Hz, 2.8 Hz, 1H), 7.01 (dd, $J = 16.4$ Hz, 0.8 Hz, 1H), 6.86 (d, $J = 2.8$ Hz, 1H), 3.12 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 180.2, 151.2, 150.4, 137.6, 133.5, 130.5, 127.2, 127.0, 125.9, 124.2, 116.4, 113.8, 109.8, 105.8, 40.5; LRMS(ESI⁺) m/z calcd for $\text{C}_{16}\text{H}_{15}\text{F}_3\text{NO}$ $[\text{M}+\text{H}]^+$ 294.11; Found 294.1.

■ (E)-1-(4-Aminophenyl)-3-(6-(dimethylamino)naphthalen-2-yl)prop-2-en-1-one
(3a)



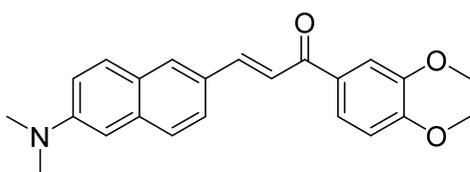
under 50 °C for 24 h. (Yield: 97%)

3a was prepared by modified general procedure **B**, *Aldol-1* procedure. Reaction was performed

$R_f = 0.05$ (1:3 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, $\text{DMSO}-d_6$) δ 8.05 (s, 1H), 7.94 (d, $J = 8.4$ Hz, 2H), 7.90 (d, $J = 8.4$ Hz, 1H), 7.85 (d, $J = 15.2$ Hz, 1H), 7.77 (d, $J = 8.8$ Hz, 1H), 7.71 (d, $J = 15.2$ Hz, 1H), 7.68 (d, $J = 8.0$ Hz, 1H), 7.24 (dd, $J = 9.2$ Hz, $J = 2.4$ Hz, 1H), 6.96 (d, $J = 2.0$ Hz, 1H), 6.63 (d, J

= 8.8 Hz, 2H), 6.12 (s, 2H), 3.04 (s, 6H); ^{13}C NMR (100 MHz, DMSO- d_6) δ 158.8, 153.7, 149.3, 142.3, 135.7, 131.0, 130.1, 129.4, 128.4, 126.4, 125.7, 125.7, 124.7, 120.0, 116.4, 112.7, 105.5, 40.1; LRMS(ESI $^+$) m/z calcd for $\text{C}_{21}\text{H}_{21}\text{N}_2\text{O}$ [M+H] $^+$ 317.17; Found 317.2.

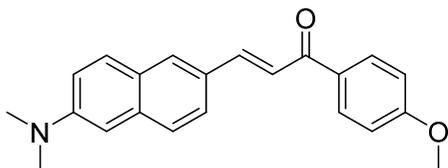
■ (*E*)-1-(3,4-Dimethoxyphenyl)-3-(6-(dimethylamino)naphthalen-2-yl)prop-2-en-1-one (**3b**)



3b was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 70%)

R_f = 0.15 (1:3 = EtOAc:n-hexane, v/v); ^1H NMR (400 MHz, CDCl_3) δ 7.95 (d, J = 16.0 Hz, 1H), 7.88 (s, 1H), 7.69–7.75 (m, 3H), 7.63–7.67 (m, 2H), 7.58 (d, J = 15.6 Hz, 1H), 7.15 (dd, J = 9.0 Hz, 2.6 Hz, 1H), 6.94 (d, J = 8.8 Hz, 1H), 6.89 (d, J = 2.8 Hz, 1H), 3.99 (s, 3H), 3.97 (s, 3H), 3.10 (s, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 188.8, 153.1, 149.7, 149.3, 145.1, 136.4, 131.9, 131.0, 129.9, 128.7, 126.9, 126.3, 124.3, 122.9, 119.4, 116.4, 110.9, 110.1, 106.0, 56.2, 56.2, 40.7; LRMS(ESI $^+$) m/z calcd for $\text{C}_{23}\text{H}_{24}\text{NO}_3$ [M+H] $^+$ 362.18; Found 362.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(4-methoxyphenyl)prop-2-en-1-one (**3c**)



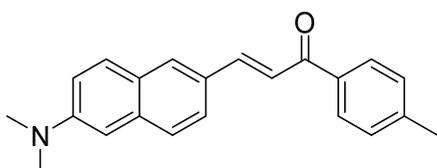
3c was prepared by general procedure

B, *Aldol-1* procedure. (Yield: 96%)

$R_f = 0.28$ (1:3 = EtOAc:n-hexane,

v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.07 (d, $J = 8.4$ Hz, 1H), 7.94 (d, $J = 15.6$ Hz, 1H), 7.88 (s, 1H), 7.54–7.75 (m, 4H), 7.15 (dd, $J = 9.0$ Hz, 2.2 Hz, 1H), 6.99 (d, $J = 9.2$ Hz, 2H), 6.89 (d, $J = 2.4$ Hz, 1H), 3.90 (s, 3H), 3.10 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 189.0, 163.3, 149.7, 145.1, 136.4, 131.7, 130.9, 130.9, 129.9, 128.7, 126.9, 126.3, 124.4, 119.7, 116.5, 113.9, 106.1, 55.6, 40.7; LRMS(ESI⁺) m/z calcd for $\text{C}_{22}\text{H}_{22}\text{NO}_2$ [M+H]⁺ 332.17; Found 332.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(p-tolyl)prop-2-en-1-one (**3d**)



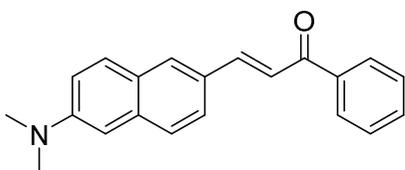
3d was prepared by general procedure

B, *Aldol-1* procedure. (Yield: 48%)

$R_f = 0.48$ (1:3 = EtOAc:n-hexane, v/v);

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.97 (d, $J = 8.0$ Hz, 1H), 7.90 (d, $J = 18.4$ Hz, 1H), 7.61–7.75 (m, 3H), 7.55 (d, $J = 15.6$ Hz, 1H), 7.30 (d, $J = 8.4$ Hz, 2H), 7.15 (dd, $J = 9.2$ Hz, 2.4 Hz, 1H), 6.88 (d, $J = 2.0$ Hz, 1H), 3.09 (s, 6H), 2.44 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 190.3, 149.7, 145.5, 143.4, 136.5, 136.3, 131.0, 129.9, 129.4, 128.7, 128.7, 126.9, 126.3, 124.4, 119.9, 116.5, 106.0, 40.7, 21.8; LRMS(ESI⁺) m/z calcd for $\text{C}_{22}\text{H}_{22}\text{NO}$ [M+H]⁺ 316.17; Found 316.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-phenylprop-2-en-1-one (**3e**)

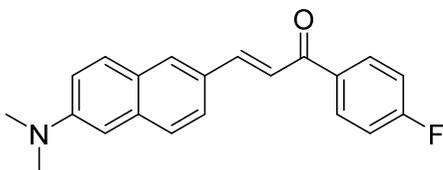


3e was prepared by general procedure **B**,
Aldol-1 procedure. (Yield: 88%)

$R_f = 0.45$ (1:3 = EtOAc:n-hexane, v/v);

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.06 (dd, $J = 8.2$ Hz, 1.4 Hz, 1H), 7.96 (d, $J = 15.6$ Hz, 1H), 7.88 (s, 1H), 7.48–7.75 (m, 7H), 7.15 (dd, $J = 9.0$ Hz, 2.6 Hz, 1H), 6.88 (d, $J = 2.8$ Hz, 1H), 3.09 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 190.7, 149.7, 146.0, 138.8, 136.5, 132.6, 131.2, 129.9, 128.7, 128.6, 128.5, 126.9, 126.2, 124.3, 119.8, 116.4, 105.9, 40.6; LRMS(ESI⁺) m/z calcd for $\text{C}_{21}\text{H}_{20}\text{NO}$ [M+H]⁺ 302.15; Found 302.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(4-fluorophenyl)prop-2-en-1-one (**3f**)

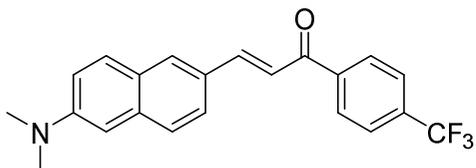


3f was prepared by general procedure
B, *Aldol-1* procedure. (Yield: 46%)

$R_f = 0.45$ (1:3 = EtOAc:n-hexane,

v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.06–8.11 (m, 2H), 7.95 (d, $J = 16.0$ Hz, 1H), 7.88 (s, 1H), 7.63–7.76 (m, 3H), 7.51 (d, $J = 15.6$ Hz, 1H), 7.12–7.22 (m, 3H), 6.88 (d, $J = 1.2$ Hz, 1H), 3.10 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 189.0, 149.8, 146.2, 136.6, 131.3, 131.1, 131.1, 130.0, 128.4, 127.0, 126.2, 124.3, 119.3, 116.5, 115.9, 115.6, 106.0, 40.6; LRMS(ESI⁺) m/z calcd for $\text{C}_{21}\text{H}_{19}\text{FNO}$ [M+H]⁺ 320.15; Found 320.1.

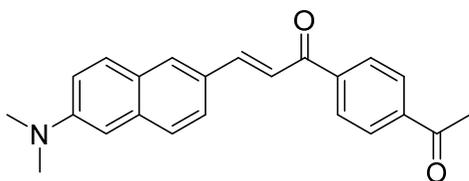
- **(E)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(4-(trifluoromethyl)phenyl)prop-2-en-1-one (3g)**



3g was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 80%)

$R_f = 0.76$ (1:1 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.12 (d, $J = 7.6$ Hz, 2H), 7.96 (d, $J = 16.0$ Hz, 1H), 7.89 (s, 1H), 7.63–7.78 (m, 5H), 7.49 (d, $J = 15.2$ Hz, 1H), 7.15 (dd, $J = 8.8$ Hz, 2.8 Hz, 2H), 6.88 (d, $J = 2.8$ Hz, 1H), 3.11 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 189.9, 149.9, 147.3, 141.8, 136.8, 131.7, 130.1, 128.8, 128.1, 127.1, 126.2, 125.7, 125.7, 124.2, 123.7, 119.3, 116.5, 105.9, 40.6; LRMS(ESI $^+$) m/z calcd for $\text{C}_{22}\text{H}_{19}\text{F}_3\text{NO}$ [M+H] $^+$ 370.14; Found 370.1.

- **(E)-1-(4-Acetylphenyl)-3-(6-(dimethylamino)naphthalen-2-yl)prop-2-en-1-one (3h)**



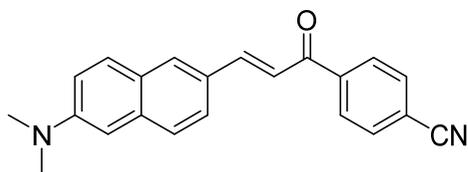
3h was prepared by modified general procedure **B**, *Aldol-1* procedure. 5 equiv. of reacting substrate, 1,4-diacetylbenzene, was used. (Yield: 23%)

$R_f = 0.55$ (1:1 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.09 (dd, $J = 12.0$ Hz, 8.4 Hz, 4H), 7.96 (d, $J = 15.6$ Hz, 1H), 7.89 (s, 1H), 7.62–7.76 (m, 3H), 7.51 (d, $J = 15.6$ Hz, 1H), 7.15 (dd, $J = 8.8$ Hz, 2.8 Hz, 1H), 6.88 (d, $J = 2.4$ Hz, 1H), 3.10 (s, 6H), 2.67 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 197.8,

190.2, 149.9, 147.0, 142.4, 139.7, 136.8, 131.6, 130.0, 128.7, 128.6, 128.2, 127.0, 126.2, 124.3, 119.5, 116.5, 105.9, 40.6, 27.1; LRMS(ESI⁺) m/z calcd for C₂₃H₂₂NO₂ [M+H]⁺ 344.17; Found 344.2.

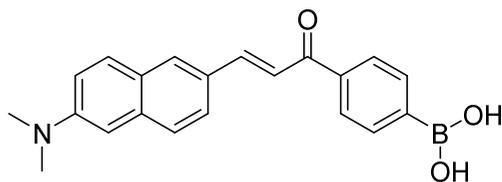
■ (*E*)-4-(3-(6-(Dimethylamino)naphthalen-2-yl)acryloyl)benzonitrile (**3i**)

3i was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 64%)



$R_f = 0.68$ (1:1 = EtOAc:n-hexane, v/v); ¹H NMR (400 MHz, CDCl₃) δ 8.09 (d, $J = 8.4$ Hz, 2H), 7.96 (d, $J = 16.0$ Hz, 1H), 7.88 (s, 1H), 7.79 (d, $J = 8.4$ Hz, 2H), 7.62–7.75 (m, 3H), 7.45 (d, $J = 15.6$ Hz, 1H), 7.15 (dd, $J = 8.8$ Hz, 2.4 Hz, 1H), 6.87 (d, $J = 1.6$ Hz, 1H), 3.11 (s, 6H); ¹³C NMR (100 MHz, CDCl₃) δ 189.2, 150.0, 147.7, 142.3, 136.9, 132.5, 131.9, 130.1, 128.9, 127.9, 127.1, 126.1, 124.2, 118.8, 118.3, 116.5, 115.7, 105.9, 40.6; LRMS(ESI⁺) m/z calcd for C₂₂H₁₉N₂O [M+H]⁺ 327.15; Found 327.2.

■ (*E*)-(4-(3-(6-(Dimethylamino)naphthalen-2-yl)acryloyl)phenyl)boronic acid (**3j**)

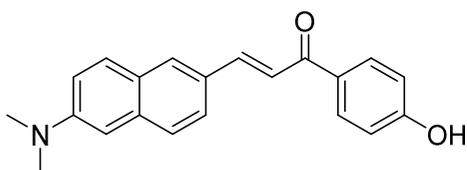


3j was prepared by modified general procedure **B**, *Aldol-1* procedure. Reaction was performed under 50 °C for 24 h. (Yield: 64%)

$R_f = 0.24$ (1:20 = methanol:DCM, v/v); ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.30 (s, 2H), 8.13 (s, 1H), 8.12 (d, $J = 7.6$ Hz, 2H), 7.91–7.98 (m, 4H), 7.84 (d, $J =$

15.2 Hz, 1H), 7.79 (d, $J = 9.2$ Hz, 1H), 7.70 (d, $J = 8.8$ Hz, 1H), 7.25 (dd, $J = 9.0$ Hz, $J = 2.2$ Hz, 1H), 6.97 (d, $J = 2.0$ Hz, 1H), 3.06 (s, 6H); ^{13}C NMR (100 MHz, DMSO- d_6) δ 189.1, 149.5, 145.0, 139.1, 136.1, 134.3, 131.1, 129.7, 127.9, 127.2, 126.5, 125.6, 124.8, 119.6, 116.4, 105.4, 40.1; LRMS(ESI $^+$) m/z calcd for $\text{C}_{21}\text{H}_{21}\text{BNO}_3$ [M+H] $^+$ 346.16; Found 346.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(4-hydroxyphenyl)prop-2-en-1-one (**3k**)

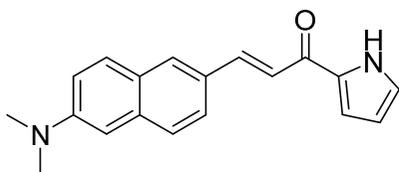


3k was prepared by modified general procedure **B**, *Aldol-1* procedure. Reaction was performed

under 70 °C for 12 h. (Yield: 33%)

$R_f = 0.27$ (1:3 = EtOAc:n-hexane, v/v); ^1H NMR (400 MHz, DMSO- d_6) δ 8.06–8.10 (m, 3H), 7.90–7.94 (m, 1H), 7.87 (s, 1H), 7.80 (s, 1H), 7.76 (d, $J = 4.0$ Hz, 1H), 7.69 (d, $J = 8.8$ Hz, 1H), 7.25 (dd, $J = 9.4$ Hz, 2.6 Hz, 1H), 6.96 (d, $J = 2.4$ Hz, 1H), 6.90 (d, $J = 8.4$ Hz, 2H), 3.34 (brs, 1H), 3.05 (s, 6H); ^{13}C NMR (100 MHz, DMSO- d_6) δ 187.0, 162.1, 149.4, 143.6, 135.9, 131.0, 130.6, 129.5, 129.4, 128.2, 126.5, 125.6, 124.7, 119.6, 116.4, 115.4, 105.4, 40.1; LRMS(ESI $^+$) m/z calcd for $\text{C}_{21}\text{H}_{20}\text{NO}_2$ [M+H] $^+$ 318.15; Found 318.1.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(1H-pyrrol-2-yl)prop-2-en-1-one (**4a**)

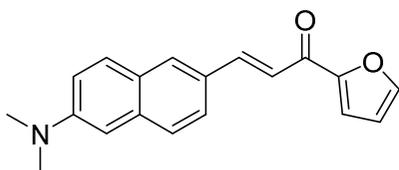


(Yield: 69%)

4a was prepared by modified general procedure **B**, *Aldol-1* procedure. Reaction was performed under 60 °C for 86 h.

R_f = 0.19 (1:3 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 9.73 (brs, 1H), 7.96 (d, J = 15.6 Hz, 1H), 7.88 (s, 1H), 7.74 (d, J = 8.81 Hz, 1H), 7.70 (dd, J = 8.6 Hz, J = 1.4 Hz, 1H), 7.64 (d, J = 8.8 Hz, 1H), 7.38 (d, J = 15.6 Hz, 1H), 7.31 (dd, J = 9 Hz, J = 2.6 Hz, 1H), 7.10 (m, 2H), 6.89 (d, J = 2.4 Hz, 1H), 6.72 (m, 1H), 3.10 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 179.2, 149.7, 143.3, 136.4, 133.6, 130.7, 129.8, 128.7, 126.9, 126.3, 124.9, 124.4, 119.7, 116.5, 115.8, 111.0, 106.1, 40.7; LRMS(ESI $^+$) m/z calcd for $\text{C}_{19}\text{H}_{19}\text{N}_2\text{O}$ $[\text{M}+\text{H}]^+$ 291.15; Found 291.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(furan-2-yl)prop-2-en-1-one
(**4b**)

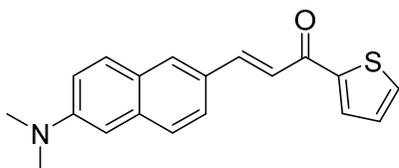


4b was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 94%)

R_f = 0.31 (1:3 = EtOAc:n-hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 8.01 (d, J = 15.6 Hz, 1H), 7.89 (s, 1H), 7.74 (d, J = 9.2 Hz, 1H), 7.70 (d, J = 9.6 Hz, 1H), 7.64 (d, J = 9.6 Hz, 1H), 7.47 (d, J = 15.6 Hz, 1H), 7.33 (d, J = 3.2 Hz, 1H), 7.15 (dd, J = 9.6 Hz, J = 2.4 Hz, 1H), 6.88 (d, J = 2.4 Hz, 1H), 6.60 (dd, J = 3.2 Hz, J = 1.2 Hz, 1H), 3.10 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 178.3, 154.2, 149.8, 146.3, 145.1, 136.6, 131.3, 130.0, 128.4, 126.9, 126.2, 124.4, 118.9, 117.1, 116.4, 112.6, 106.0, 40.7;

LRMS(ESI⁺) m/z calcd for C₁₉H₁₈NO₂ [M+H]⁺ 292.13; Found 292.2.

- (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(thiophen-2-yl)prop-2-en-1-one (**4c**)



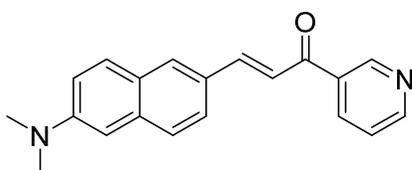
4c was prepared by general procedure **B**,

Aldol-1 procedure. (Yield: 77%)

R_f = 0.18 (1:5 = EtOAc:n-hexane, v/v);

¹H NMR (400 MHz, CDCl₃) δ 7.99 (d, J = 15.2 Hz, 1H), 7.88–7.91 (m, 2H), 7.62–7.76 (m, 4H), 7.44 (d, J = 16.0 Hz, 1H), 7.13–7.21 (m, 2H), 6.88 (d, J = 2.4 Hz, 1H), 3.10 (s, 6H); ¹³C NMR (100 MHz, CDCl₃) δ 182.3, 149.8, 146.2, 145.2, 136.6, 133.5, 131.5, 131.3, 130.0, 128.3, 128.3, 127.0, 126.2, 124.3, 119.3, 116.5, 106.0, 40.7; LRMS(ESI⁺) m/z calcd for C₁₉H₁₈NOS [M+H]⁺ 308.11; Found 308.1.

- (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(pyridin-3-yl)prop-2-en-1-one (**4d**)



4d was prepared by general procedure **B**,

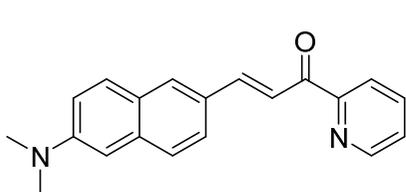
Aldol-1 procedure. (Yield: 54%)

R_f = 0.25 (1:1 = EtOAc:n-hexane, v/v);

¹H NMR (500 MHz, CDCl₃) δ 9.26 (d, J = 2.0 Hz, 1H), 8.79 (dd, J = 5.0 Hz, 1.5 Hz, 1H), 8.30 (d, J = 8.0 Hz, 1H), 7.97 (d, J = 16.0 Hz, 1H), 7.88 (s, 1H), 7.63–7.74 (m, 3H), 7.49 (d, J = 15.5 Hz, 1H), 7.43–7.47 (m, 1H), 7.14 (d, J = 8.5 Hz, 2.5 Hz, 1H), 6.87 (d, J = 2.5 Hz, 1H), 3.10 (s, 6H); ¹³C NMR (125 MHz, CDCl₃) δ 189.2, 153.0, 149.9, 149.8, 147.1, 136.8, 136.0, 134.1, 131.8, 130.0,

128.0, 127.1, 126.1, 124.2, 123.7, 119.0, 116.4, 105.9, 40.6; LRMS(ESI⁺) m/z calcd for C₂₀H₁₉N₂O [M+H]⁺ 303.15; Found 303.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(pyridin-2-yl)prop-2-en-1-one



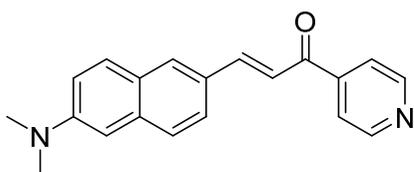
4e was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 63%)

$R_f = 0.71$ (1:1 = EtOAc:n-hexane, v/v);

¹H NMR (400 MHz, CDCl₃) δ 8.75–8.78 (m, 1H), 8.31 (d, $J = 16.0$ Hz, 1H), 8.21 (dt, $J = 7.6$ Hz, 1.1 Hz, 1H), 8.09 (d, $J = 16.4$ Hz, 1H), 7.96 (d, $J = 1.2$ Hz, 1H), 7.87 (td, $J = 7.6$ Hz, 1.7 Hz, 1H), 7.81 (dd, $J = 8.8$ Hz, 2.0 Hz, 1H), 7.74 (d, $J = 8.8$ Hz, 1H), 7.64 (d, $J = 8.8$ Hz, 1H), 7.46–7.50 (m, 1H), 7.14 (dd, $J = 8.8$ Hz, 2.8 Hz, 1H), 6.88 (d, $J = 2.4$ Hz, 1H), 3.09 (s, 6H); ¹³C NMR (100 MHz, CDCl₃) δ 189.5, 154.8, 149.8, 148.9, 146.0, 137.1, 136.6, 131.5, 130.0, 128.9, 126.8, 126.8, 126.2, 124.9, 123.0, 118.6, 116.4, 106.0, 40.7; LRMS(ESI⁺) m/z calcd for C₂₀H₁₉N₂O [M+H]⁺ 303.15; Found 303.2.

■ (*E*)-3-(6-(Dimethylamino)naphthalen-2-yl)-1-(pyridin-4-yl)prop-2-en-1-one

(4f)



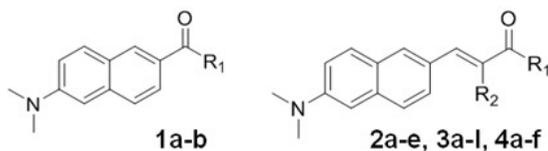
4f was prepared by general procedure **B**, *Aldol-1* procedure. (Yield: 42%)

$R_f = 0.29$ (1:1 = EtOAc:n-hexane, v/v);

¹H NMR (400 MHz, CDCl₃) δ 8.81–8.84 (m, 2H), 7.95 (d, $J = 16.0$ Hz, 1H),

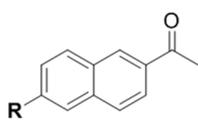
7.88 (s, 1H), 7.77–7.80 (m, 2H), 7.73 (d, $J = 9.2$ Hz, 1H), 7.62–7.69 (m, 2H), 7.42 (d, $J = 15.6$ Hz, 1H), 7.15 (dd, $J = 8.8$ Hz, 2.8 Hz, 1H), 6.87 (d, $J = 2.0$ Hz, 1H), 3.10 (s, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 189.9, 150.8, 150.0, 148.0, 145.2, 136.9, 132.0, 130.1, 127.9, 127.1, 126.1, 124.2, 121.7, 118.8, 116.4, 105.9, 40.6; LRMS(ESI⁺) m/z calcd for $\text{C}_{20}\text{H}_{19}\text{N}_2\text{O}$ $[\text{M}+\text{H}]^+$ 303.15; Found 303.2.

1.1.6. Supporting Information



Cpd.	R ₁	R ₂	λ _{exc}	λ _{em}	E _{HOMO}	E _{LUMO}	ΔE (H-L)	ΔE (S ₀ -S ₁)	f
1a	CH ₃		354	455	-0.1886	-0.0487	0.1399	3.4151	0.292
1b	H		373	463	-0.1925	-0.0568	0.1357	3.5002	0.293
2a	N(CH ₃) ₂	H	361	485	-0.1808	-0.0529	0.1279	3.2195	0.541
2b	OCH ₃	H	382	507	-0.1853	-0.0607	0.1246	3.1613	0.553
2c	CH ₃	H	388	536	-0.1860	-0.0608	0.1252	3.2388	0.707
2d	OCH ₃	CN	423	591	-0.1960	-0.0803	0.1157	3.0045	0.771
2e	CF ₃	H	461	632	-0.1957	-0.0862	0.1096	2.8436	0.595
3a	4-NH ₂ Ph	H	414	565	-0.1796	-0.0607	0.1190	2.9695	0.784
3b	3,4-(CH ₃ O) ₂ Ph	H	422	584	-0.1837	-0.0667	0.1171	2.9276	0.735
3c	4-CH ₃ OPh	H	423	586	-0.1832	-0.0671	0.1161	2.9177	0.719
3d	4-CH ₃ Ph	H	420	590	-0.1846	-0.0696	0.1149	2.8948	0.666
3e	Ph	H	423	599	-0.1856	-0.0712	0.1144	2.8857	0.643
3f	4-FPh	H	423	602	-0.1875	-0.0741	0.1134	2.8642	0.659
3g	4-CF ₃ Ph	H	436	650	-0.1908	-0.0821	0.1087	2.7474	0.617
3h	4-CH ₃ COPh	H	444	657	-0.1889	-0.0835	0.1055	2.6398	0.515
3i	4-CNPh	H	449	667	-0.1929	-0.0888	0.1040	2.6139	0.568
4a	pyrrol-2-yl	H	412	567	-0.1852	-0.0667	0.1186	2.9751	0.704
4b	furan-2-yl	H	427	601	-0.1841	-0.0710	0.1131	2.8525	0.658
4c	thiophen-2-yl	H	431	612	-0.1854	-0.0735	0.1119	2.8083	0.606
4d	pyrid-3-yl	H	435	625	-0.1895	-0.0778	0.1117	2.8266	0.627
4e	pyrid-2-yl	H	437	633	-0.1830	-0.0738	0.1092	2.7588	0.596
4f	pyrid-4-yl	H	441	646	-0.1907	-0.0821	0.1086	2.7522	0.601
3j	4-boronophenyl	H	424	598	-0.1895	-0.0780	0.1115	2.8291	0.741
3k	4-hydroxyphenyl	H	416	575	-0.1841	-0.0683	0.1158	2.912	0.686

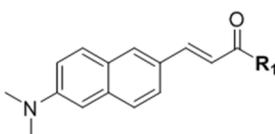
Table S1.1.1. Photophysical properties of all naphthalene-based two-photon fluorophores with chemical structures. Observed excitation and emission wavelength (nm), calculated HOMO and LUMO energy levels (eV), calculated S₀-S₁ energy gaps (Hartree) and oscillator strengths of compounds **1a-b**, **2a-e**, **3a-i**, and **4a-f**.



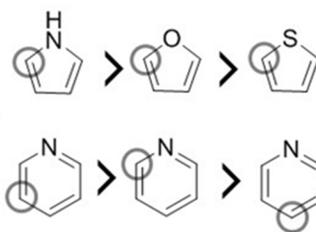
λ_{ex} : S_0 to S_1 excitation wavelength (nm)
 f : oscillator strength

R	λ_{ex} (nm)	f	R	λ_{ex} (nm)	f
 (1a)	354	0.292		322	0.071
	342	0.261		336	0.226
	339	0.221		328	0.203
	329	0.175		296	0.110
				285	0.065

Figure S1.1.1. Predicted S_0 to S_1 excitation wavelength and oscillator strength of dimethylamino-containing naphthalene-based analogues from **1a**.



Cpd.	R_1	λ_{em} (nm)
4a	pyrrol-2-yl	567
4b	furan-2-yl	601
4c	thiophen-2-yl	612
4d	pyrid-3-yl	625
4e	pyrid-2-yl	633
4f	pyrid-4-yl	646
3e	Ph	599



Figures S1.1.2. Electron-donating ability of heteroaromatic ring systems could be estimated by emission wavelength (λ_{em} , nm) and HOMO-LUMO energy gaps (eV); Increasing their electron-donating ability of heteroaromatic rings or their substituted positions were resulted in lowering their emission wavelengths.

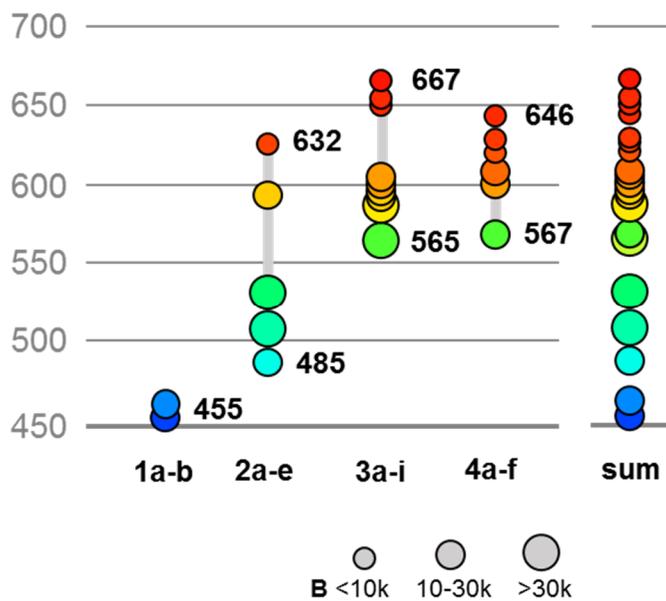
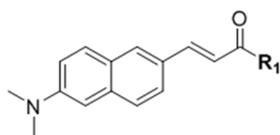


Figure S1.1.3. Brightness in one-photon excited fluorescence ($\epsilon\Phi$) and emission wavelength (nm) of **1a-b**, **2a-e**, **3a-i**, and **4a-f**; the size of circle represents brightness of fluorophore; y-value and color of circle represents the emission wavelength of each fluorophore.

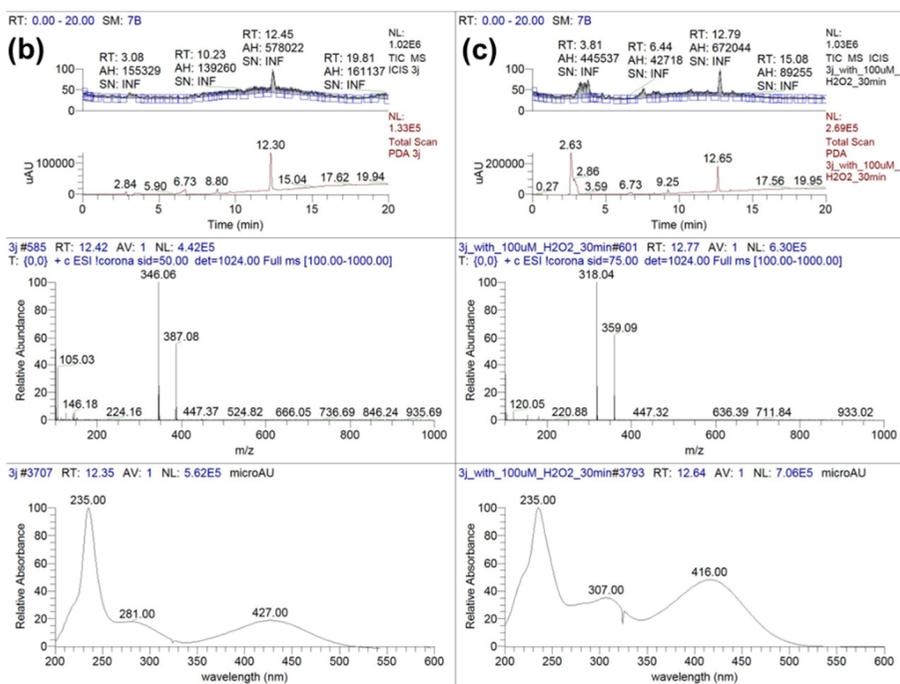
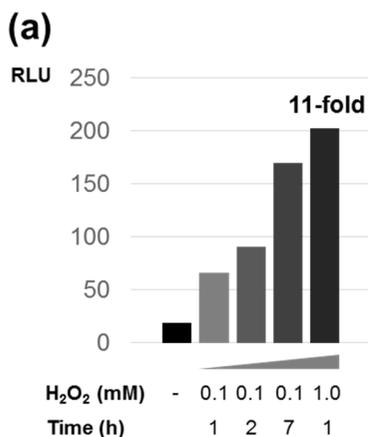


Cpd.	R ₁	λ _{em} (nm)	1/λ _{em} (μm ⁻¹)	σ _p
3a	4-NH ₂ Ph	565	1.786	-0.660
3b	3,4-(CH ₃ O) ₂ Ph	584	1.712	
3c	4-CH ₃ OPh	586	1.706	-0.268
3d	4-CH ₃ Ph	590	1.695	-0.170
3e	Ph	599	1.669	0.000
3f	4-FPh	602	1.661	0.062
3g	4-CF ₃ Ph	650	1.538	0.540
3h	4-CH ₃ COPh	657	1.522	0.502
3i	4-CNPh	667	1.499	0.660

Figure S1.1.4. General tendency between Hammett constant (σ) and inverse of fluorescence emission wavelength (nm) in **3a–i**.

Cpd.	R ₁	R ₂	λ _{ex}	λ _{em}	δΦ	λ _{ex}	ε	Φ	εΦ
3j	4-boronophenyl	H	880	598	13	424	22k	0.37	8.1k
3k	4-hydroxyphenyl	H	820	575	71	416	33k	0.45	15k

Figures S1.1.5. Photophysical properties of compounds **3j** and **3k**; λ_{ex} = fluorescence excitation wavelength (nm); λ_{em} = emission wavelength (nm); δΦ = two-photon cross section (GM, 10⁻⁵⁰ cm⁴ · s · photon⁻¹); ε = molar extinction coefficient (L · mol⁻¹ · cm⁻¹); properties were measured in DMF.



Figures S1.1.6. (a) Dose- and time-dependent increases of red fluorescence at 641 nm. LC-MS spectra of (b) **3j** and (c) **3j** with 100 μ M H₂O₂ treatment for 30 min. **3j** was prepared for 10 μ M ACN/PBS(1:1) solution in both experiments.

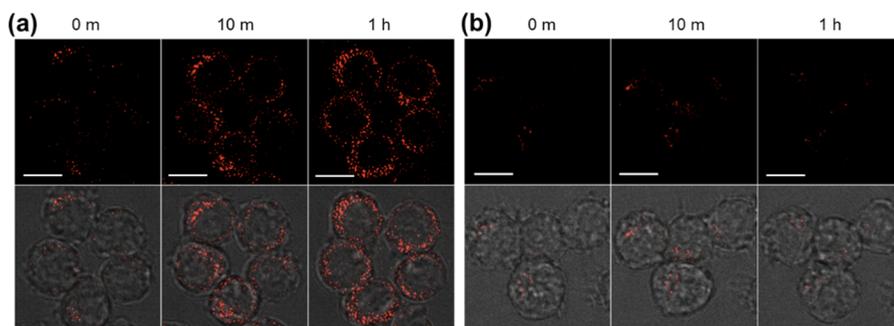


Figure S1.1.7. Fluorescence images of Raw264.7 macrophage cells pretreated with **3j** (1 μ M, 20 min) detecting endogenous H_2O_2 in a time-dependent manner. Cells were pretreated with 1 μ M **3j** for 20 min, and then time-lapse images were taken upon addition of (a) LPS (500 ng/mL) or (b) vehicle. Fluorescence images (red) were taken from Ex 438 nm/Em 632 nm. Scale bar is 5 μ m.

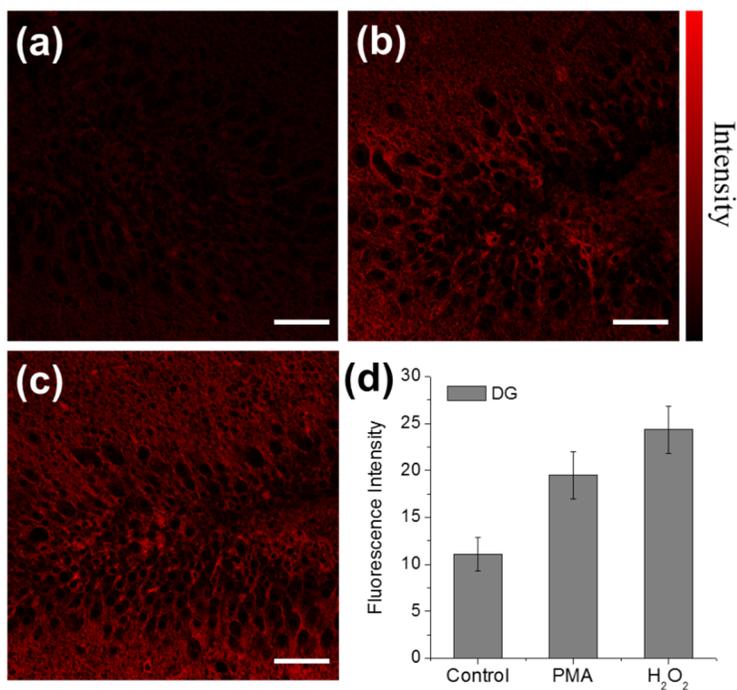


Figure S1.1.8. TPM images of a rat hippocampal slice in the DG region stained with (a) 5 μ M

3j. A rat hippocampal slices were pretreated with (b) PMA ($5 \mu\text{g mL}^{-1}$) for 30 min and (c) 1 mM H_2O_2 for 30 min before labeling with $5 \mu\text{M}$ **3j**. (d) Average TPEF intensities in (a–c). The TPEF were collected at 550–700 nm upon excitation at 820 nm with fs pulse. Scale bars (a–c) are $48 \mu\text{m}$.

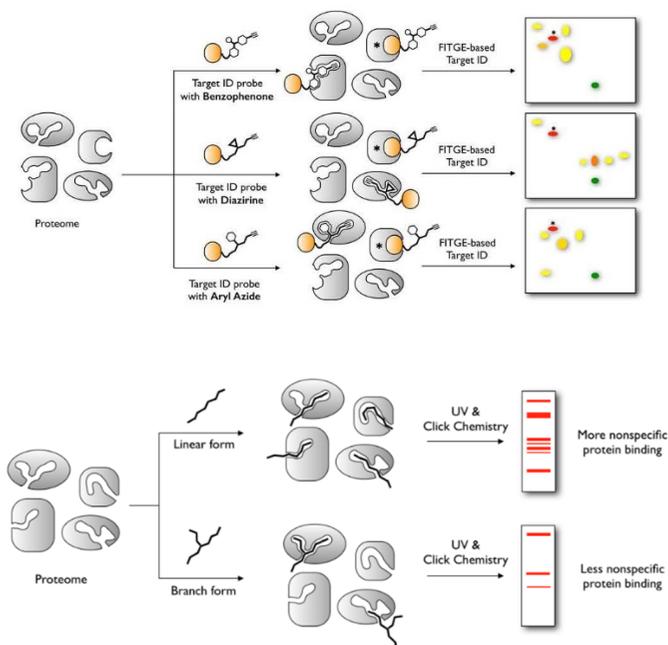
1.1.7. References and Notes

- † Portions of this chapter have been previously reported, see: Koo, J. Y.; Heo, C. H.; Shin, Y.; Kim, D.; Lim, C. S.; Cho, B. R.; Kim, H. M.; Park, S. B. *Chem. Eur. J.* **2016**, *22*, 14166–14170.
- (1) (a) Frangioni, J. V. *Curr. Opin. Chem. Biol.* **2003**, *7*, 626–634. (b) Licha, K.; Olbrich, C. *Adv. Drug Delivery Rev.* **2005**, *57*, 1087–1108. (c) Ntziachristos, V. *Annu. Rev. Biomed. Eng.* **2006**, *8*, 1–33. (d) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer-Verlag: New York, 2006.
- (2) (a) Zipfel, W. R.; Williams, R. M.; Webb, W. W. *Nat. Biotechnol.* **2003**, *21*, 1369–1377. (b) Helmchen, F.; Denk, W. *Nat. Methods* **2005**, *2*, 932–940. (c) Oheim, M.; Michael, D. J.; Geisbauer, M.; Madsen, D.; Chow, R. H. *Adv. Drug Delivery Rev.* **2006**, *58*, 788–808. (d) Rumi, M.; Perry, J. W. *Adv. Opt. Photonics* **2010**, *2*, 451–518.
- (3) (a) Kim, H. M.; Cho, B. R. *Chem. Commun.* **2009**, 153–164. (b) Kim, H. M.; Cho, B. R. *Acc. Chem. Res.* **2009**, *42*, 863–872. (c) Kim, H. M.; Cho, B. R. *Chem. Rev.* **2015**, *115*, 5014–5055.
- (4) Singha, S.; Kim, D.; Roy, B.; Sambasivan, S.; Moon, H.; Rao, A. S.; Kim, J. Y.; Joo, T.; Park, J. W.; Rhee, Y. M.; Wang, T.; Kim, K. H.;

- Shin, Y. H.; Jung, J.; Ahn, K. H. *Chem. Sci.* **2015**, *6*, 4335–4342.
- (5) (a) Kim, M. K.; Lim, C. S.; Hong, J. T.; Han, J. H.; Jang, H.-Y.; Kim, H. M.; Cho, B. R. *Angew. Chem. Int. Ed.* **2010**, *49*, 364–367. (b) Kim, H. J.; Han, J. H.; Kim, M. K.; Lim, C. S.; Kim, H. M.; Cho, B. R. *Angew. Chem. Int. Ed.* **2010**, *49*, 6786–6789. (c) Mohan, P. S.; Lim, C. S.; Tian, Y. S.; Roh, W. Y.; Lee, J. H.; Cho, B. R. *Chem. Commun.* **2009**, 5365–5367.
- (6) Kim, H. M.; An, M. J.; Hong, J. H.; Jeong, B. H.; Kwon, O.; Hyon, J.-Y.; Hong, S.-C.; Lee, K. J.; Cho, B. R. *Angew. Chem., Int. Ed.* **2008**, *47*, 2231–2234.
- (7) Chung, C.; Srikun, D.; Lim, C. S.; Chang, C. J.; Cho, B. R. *Chem. Commun.*, **2011**, *47*, 9618–9620.
- (8) (a) Lee, J. H.; Lim, C. S.; Tian, Y. S.; Han, J. H.; Cho, B. R. *J. Am. Chem. Soc.* **2010**, *132*, 1216–1217. (b) Kim, H. M.; Jeong, B. H.; Hyon, J.-Y.; An, M. J.; Seo, M. S.; Hong, J. H.; Lee, K. J.; Kim, C. H.; Joo, T.; Hong, S.-C.; Cho, B. R. *J. Am. Chem. Soc.* **2008**, *130*, 4246–4247. (c) Cohen, B. E.; McAnaney, T. B.; Park, E. S.; Jan, Y. N.; Boxer, S. G.; Jan, L. Y. *Science* **2002**, *296*, 1700–1703.
- (9) Silva, A. P.; Gunaratne, H. Q. N.; Gunnlaugsson, T.; Huxley, A. J. M.; McCoy, C. P.; Rademacher, J. T.; Rice, T. E. *Chem. Rev.* **1997**, *97*, 1515–1566.
- (10) (a) Zoon, P. D.; van Stokkum, I. H. M.; Parent, M.; Mongin, O.; Blanchard-Desce, M.; Brouwer, A. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2706–2715. (b) Kim, E.; Koh, M.; Ryu, J.; Park, S. B. *J. Am. Chem.*

- Soc.* **2008**, *130*, 12206–12207. (c) Kim, E.; Koh, M.; Lim, B. J.; Park, S. B. *J. Am. Chem. Soc.* **2011**, *133*, 6642–6649.
- (11) (a) Murugan, N. A.; Kongsted, J.; Ågren, H. *J. Chem. Theory Comput.* **2013**, *9*, 3660–3669. (b) Kim, E.; Lee, Y.; Lee, S.; Park, S. B. *Acc. Chem. Res.* **2015**, *48*, 538–547.
- (12) (a) Hilborn, R. C. *Am. J. Phys.* **1982**, *50*, 982–986. (b) Lee, Y.; Jo, A.; Park, S. B. *Angew. Chem., Int. Ed.* **2015**, *54*, 15689–15693.
- (13) Rao, A. S.; Kim, D.; Wang, T.; Kim, K. H.; Hwang, S.; Ahn, K. H. *Org. Lett.* **2012**, *14*, 2598–2601.
- (14) Chang, C.-E.; Gilson, M. K. *J. Comput. Chem.* **2003**, *24*, 1987–1998.
- (15) Lee, S. K.; Yang, W. J.; Choi, J. J.; Kim, C. H.; Jeon, S.-J.; Cho, B. R. *Org. Lett.* **2005**, *7*, 323–326.
- (16) Makarov, N. S.; Drobizhev, M.; Rebane, A. *Opt. Express* **2008**, *16*, 4029–4047.
- (17) Hyder, Z.; Ruan, J.; Xiao, J. *Chem. Eur. J.* **2008**, *14*, 5555–5566.
- (18) Rao, A. S.; Kim, D.; Wang, T.; Kim, K. H.; Hwang, S.; Ahn, K. H. *Org. Lett.* **2012**, *14*, 2598–2601.
- (19) Hirohito, U.; Kyoko, T.; Anwar; Xuan-Ming, D.; Shuji, O.; Hidetoshi, O.; Hiro, M.; Hachiro, N. *Nonlinear Optics* **2000**, *24*, 73–78.

Part 2. Design and Optimization of Target Identification Probes



2.1. Probe Design for Target Protein Identification of Bioactive Small Molecules: Design Strategy and Case Studies

2.1.1. Abstract

There is a growing demand for the development of first-in-class therapeutic agents and the discovery of a novel mode of action in various diseases. For this endeavor, phenotypic screening has been one of the major discovery-driven approaches to reveal a new class of drug candidates. However, optimization of original hit compound derived from phenotypic screening is extremely challenging due to the lack of information on their target proteins; therefore, a rational drug discovery approach is not available. For this purpose, optimization of bioactive hit compounds should be pursued along with target identification from the beginning of drug discovery process.

Among the increasing interest in target identification method, FITGE (Fluorescence difference In Two-dimensional Gel Electrophoresis) technique was discovered. FITGE method has its own advantages; covalent crosslinking between protein and ligand to find out the target proteins of low concentration or low binding affinity, and high selectivity based on LC-MS/MS analysis of 2D-gel fraction. Prior to performing FITGE-based target identification, design of a proper target ID probe is necessary. Target ID probe includes original hit compound structure as well as photoactivable crosslinker and biorthogonal

functional moiety, without the elimination of original bioactivity.

Herein, we focus on the design strategy of FITGE-based target ID probes based on two case studies; the discovery and identification of the target protein of a novel anti-inflammatory agent for treating neurodegenerative diseases and a novel glucose uptake enhancer for type 2 diabetes.

2.1.2. Introduction

In medical purpose, small molecule drugs have been used for thousands of years. Nature has always been an indispensable source of acquiring bioactive small molecules, which could be extracted from plants, animals, bacteria or fungi. In the previous two centuries, there have been approaches to construct a set of structural derivatives of natural compounds as well as newly generated synthetic small molecules.¹

Along with thousands of approaches to find out novel drug candidate molecules, elucidating their biological mechanism is a vital and significant process for chemical biology and further medical research. In order to exert biological activities of small molecules in human body, engagement between a small molecule and its biomolecular partner is necessary.² Although the research field of discovering a new bioactive chemical entities flourished, the identification (ID) of the target protein still remained as a bottleneck process.³

Among the number of interactions in biological systems, methods for elucidating interactions between proteins and small molecules are beginning to emerge in chemical biology research. In advance, for the purpose of identifying

target proteins of small molecules, the most common approach is affinity-based methods, including affinity chromatography-pull down analysis,⁴ activity- or affinity-based proteome profiling,^{5,3a} small molecule-affinity chromatography with stable isotope labeling of amino acids in cell culture (SILAC).⁶

J. Park and coworkers recently reported a novel cell-based target identification method, namely, FITGE (Fluorescence difference In Two-dimensional Gel Electrophoresis).⁷ In this method, to overcome the limitation of weak binding affinity or low abundance of target proteins, they introduced covalent anchoring of a bioactive small-molecule to cellular target proteins in live cells using a photo-crosslinker. The resolution of the in-gel protein analysis was clearly enhanced by 2D gel electrophoresis with dual-color labeling to differentiate specific target proteins from non-specific protein binders.

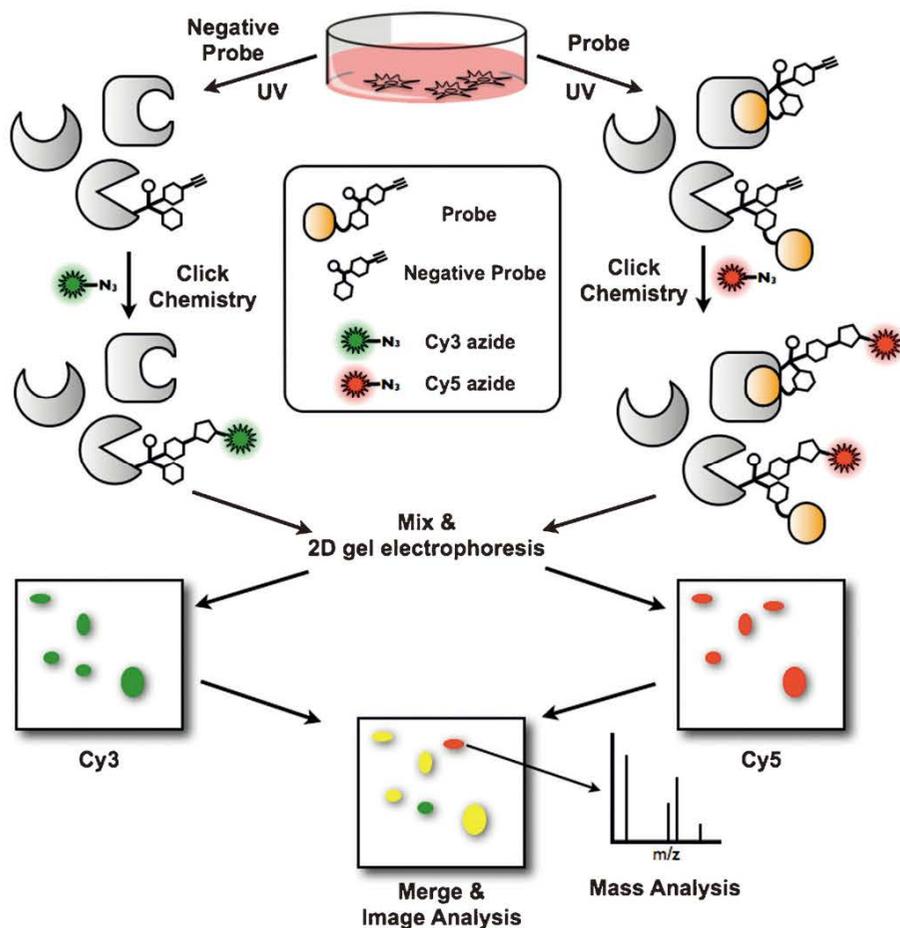
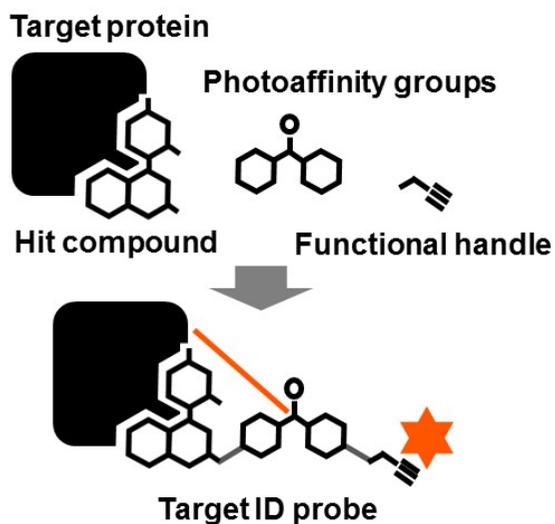


Figure 2.1.1. Outline of the FITGE method for identification of the targets of bioactive small molecules in live cells.⁷

In FITGE-based target identification, structural features of a target ID probe includes original bioactive hit compound, a photoaffinity group, and a functional handle for further modification.⁷ (Scheme 2.1.1) Besides, design of a proper chemical probe is necessary for successful target ID. First, position of additional functional groups and size of linkers are sensitive to the bioactivity of original candidate compound. Second, photoreactive groups such as

diazirine, benzophenone, and aryl azide groups can generate covalent adducts nonspecifically, because they are activated non-invasively upon irradiation with UV light, and the in situ-generated active species can create covalent bonds with nearby molecules. Selection of photoreactive groups and molecular shape of linkers are also important factors to consider. Even though there are many considerations to design a target ID probe, the paradigm shift from noncovalent to covalent photocrosslinkers has led to an increased success rate for target ID.⁸

Despite the importance of target ID probe design, there have been limited numbers of previous studies about it. Along with expanding our knowledge to pursue successful target ID, we focused on the optimization of rational strategy to design target ID probes. Herein, we introduce two case studies of target protein identification based on the rational target ID probe design. One is about an anti-neuroinflammatory agent with its identified target HMGBs, another is about a glucose uptake enhancer with its identified target PPAR γ .



Scheme 2.1.1. General strategy for target ID probe design. Target ID probe for FITGE includes original bioactive hit compound, a photoaffinity group, and a functional handle for further modification. Orange line: crosslinking between the target protein and the photoaffinity group in target ID probe; Orange figure: additional modification by copper-mediated alkyne-azide click chemistry such as attachment of fluorescent dyes or biotin.

2.1.3. Results and discussion

A Small Molecule Binding HMGB1 and HMGB2 Inhibits Microglia-Mediated Neuroinflammation – Target ID Probe Design based on Structure-activity Relationship (SAR).⁹

Microglia, a type of immune cell in the brain, serve as the first line of defense in the CNS by recognizing external pathogens, and contribute to innate and adaptive immunity by continuously surveying the microenvironment and protecting the brain. Once microglia are stimulated by immunogens or bacterial endotoxins, such as lipopolysaccharide (LPS), activated microglia play a central role in neuroinflammation and secrete various neurotoxic factors, such as IL-1 β , TNF α , PGE₂, nitric oxide (NO), and superoxide anion (O₂⁻).¹⁰ To identify anti-neuroinflammatory agents, we used a cell-based phenotypic screening with LPS-induced nitrite release in the BV-2 mouse microglial cell line as a readout. Based on a high-throughput screening of an approximately 3,500-member in-house library constructed by the pDOS strategy,¹¹ we identified a novel benzopyran-embedded tetracyclic compound, named inflachromene (ICM). ICM efficiently blocked the LPS-induced nitrite release in a dose-dependent manner without any toxicity in BV-2 microglial cells (Figure 2.1.2A). To confirm the anti-inflammatory effect of ICM in microglia, we examined the effect of ICM on other pro-inflammatory mediators. Increased levels of inflammation-related genes such as *Il6*, *Il1b*, *Nos2*, and *Tnf*, after LPS stimulation were dramatically suppressed upon treatment with ICM (Figure 2.1.2B) and the LPS-induced secretion of the pro-inflammatory cytokine

TNF was also reduced by ICM treatment. (Figure S2.1.1A) We also confirmed that ICM inhibited the nitrite release in a broad range of cell lines, including rat microglia (HAPI), mouse macrophages (RAW 264.7), and mouse primary microglial cultures (Figure S2.1.1B).

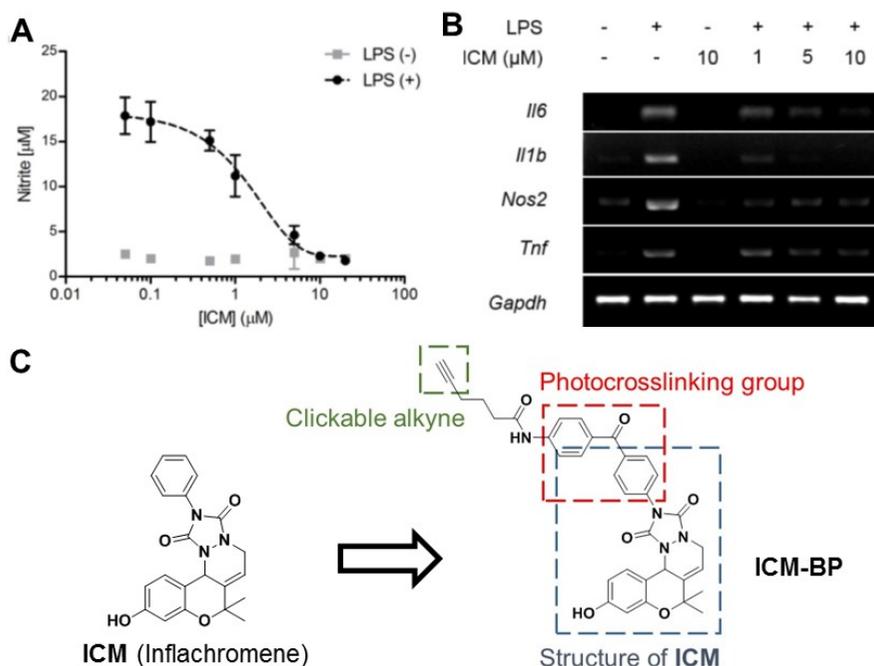


Figure 2.1.2. An anti-inflammatory agent ICM (Inflachromene) and its target ID probe design. (A) Dose-dependent inhibition of cellular nitric oxide (NO) level by an anti-inflammatory hit compound ICM. (B) RT-PCR analysis revealed the inhibitory effect of ICM on the expression levels of LPS-induced inflammation-related genes in BV-2 cells. (C) Structure of ICM and further construction of target ID probe, ICM-BP. On the basis of a structure-activity relationship study for ICM analogues (Figure 2.1.3), we designed and synthesized the chemical probe ICM-BP (Figure 2.1.2C). A benzophenone moiety was embedded as an *N*-phenyl substituent of the 1,2,4-triazoline-3,5-dione ring as a photoactivatable crosslinking moiety and an alkyne group was incorporated at the end of benzophenone-substituted linker as a functional tag for bio-orthogonal click reaction with azide-linked fluorescent dyes or biotin.

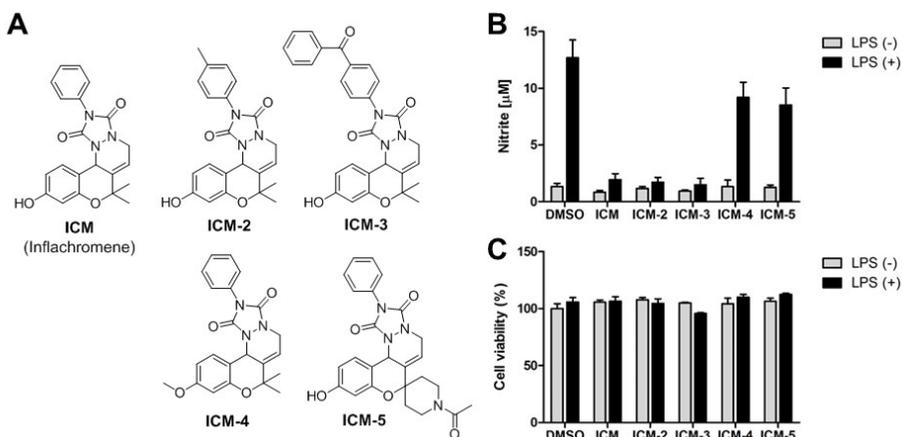


Figure 2.1.3. Structure-activity relationship study to design the chemical probe. (A) The chemical structures of ICM derivatives. (B) Anti-inflammatory activity of ICM derivatives in BV-2 cells. The inhibitory activity of ICM derivatives was estimated by the level of nitrite release by Griess assay. Cells were treated with individual derivative for 30 min prior to LPS treatment (100 ng/mL) for 24 h. (C) Non-cytotoxic effect of ICM derivatives. Cell viability was measured as described above. Data represent mean and s.d.

Prior to the application of ICM-BP for target identification, we evaluated its inhibitory activity toward LPS-induced nitrite release and confirmed that ICM-BP exhibited the inhibitory activity comparable to that of the original compound, ICM. ICM-BP also showed consistent structure-activity relationship (SAR) pattern. ICM-BP-2 exhibited low efficacy due to the position of different alkyne linker position. ICM-DA and ICM-Az which were incorporated other photoactivable crosslinkers exhibited comparable anti-inflammation efficacy with ICM or ICM-BP (Figure 2.1.4).

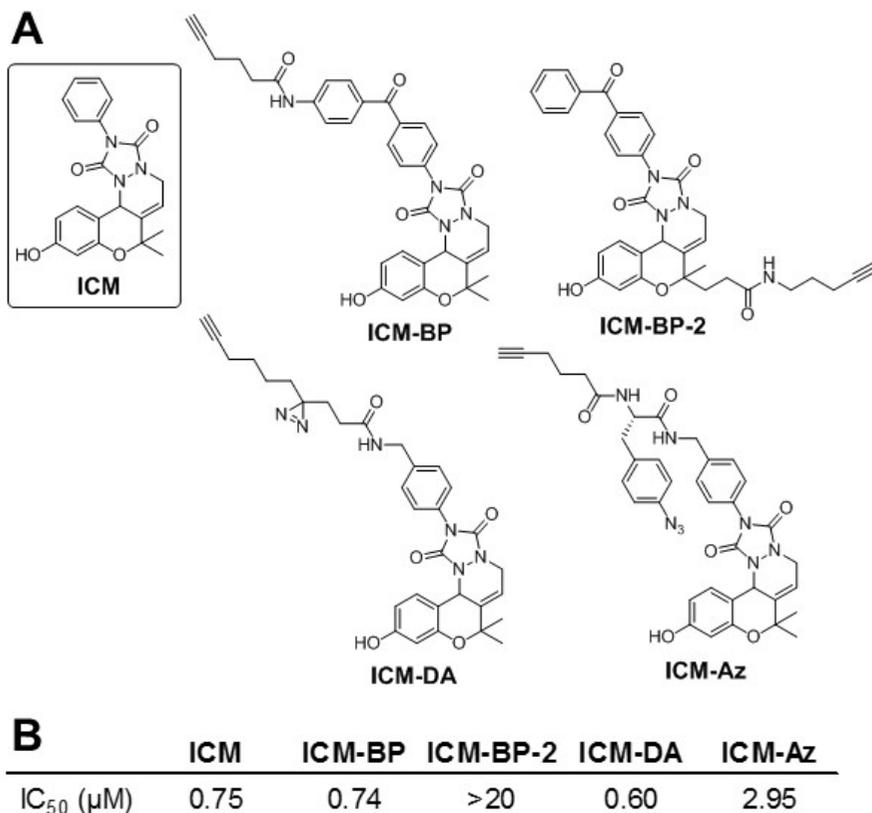


Figure 2.1.4. Structure-activity relationship of target ID probes. (A) The chemical structures of ICM and designed target ID probes of ICM. (B) anti-inflammatory IC₅₀ value of ICM and target ID probes in BV-2 cells. The inhibitory activity and IC₅₀ value of ICM and the probes were estimated by the level of nitrite release by Griess assay. Cells were treated with individual derivative for 30 min prior to LPS treatment (100 ng/mL) for 24 h.

With the optimized probe, we undertook the target identification in live cells using competitive labeling, which is an essential step for the efficient exclusion of protein bands associated with non-specific binding. Once the cells were treated with ICM-BP in the absence or presence of ICM as a soluble competitor, they were irradiated with UV light to photo-crosslink the ICM-BP to cellular

target proteins and subjected to the click reaction with azide-linked fluorescent dye. The resulting proteomes were analyzed by gel electrophoresis and gel-fluorescence scanning. As a result of 1D gel analysis, a direct comparison of the labeling pattern by ICM-BP with and without ICM competition showed two distinct bands of 29 kDa and 45 kDa (Figure 2.1.5A, denoted by arrow and asterisk) regardless of LPS stimulation. We selected the protein in the 29 kDa band, because it competed out at a lower concentration of ICM more effectively than the 45 kDa band in a dose-dependent competitive assay. The bands were excised from 1D gels of live cell labeling, and the proteins of the 29 kDa band were explored by LC-MS/MS analysis. According to MS analysis, however, numerous proteins were implicated as potential targets by 1D gel analysis (Figure 2.1.5A, right table). To narrow down the number of candidates, we performed a 2D gel analysis to improve the resolution of the in-gel analysis. ICM-BP-labeled proteomes with and without ICM competition were treated with Cy5-azide (red for specific binding) and Cy3-azide (green for non-specific binding), respectively. Two samples labeled with different fluorescent dyes were mixed and analyzed together by 2D gel electrophoresis, which reduced the gel-to-gel variation. In-gel fluorescence scanning of 2D gels effectively revealed characteristic spots around 29 kDa and 45 kDa corresponding to the bands in the 1D gel analysis (Figure 2.1.5B, 29kDa spot denoted by arrow). Following MS/MS analysis, we obtained a shorter list of proteins detected from the 29 kDa spot than in 1D gel analysis, and we selected a number of potential candidates for target proteins that were identified in both the 1D and 2D gel analyses (Figure 2.1.5A and 2.1.5B, right tables).

Among those candidate proteins, we focused on the high mobility group box 2 (HMGB2) protein as a potential target, because there are a number of reports in the literature indicating that the HMGBs play crucial roles in inflammation.¹²⁻¹⁴ To confirm the specific binding event of ICM-BP with HMGB2, we performed a pull-down assay using an avidin-based enrichment of an ICM-BP-treated proteome labeled with biotin-azide via a click reaction. This affinity pull-down assay confirmed the specific binding of ICM-BP with HMGB2, which was reinforced by the complete abolishment of the binding event of ICM-BP with HMGB2 upon ICM competition (Figure S2.1.2).

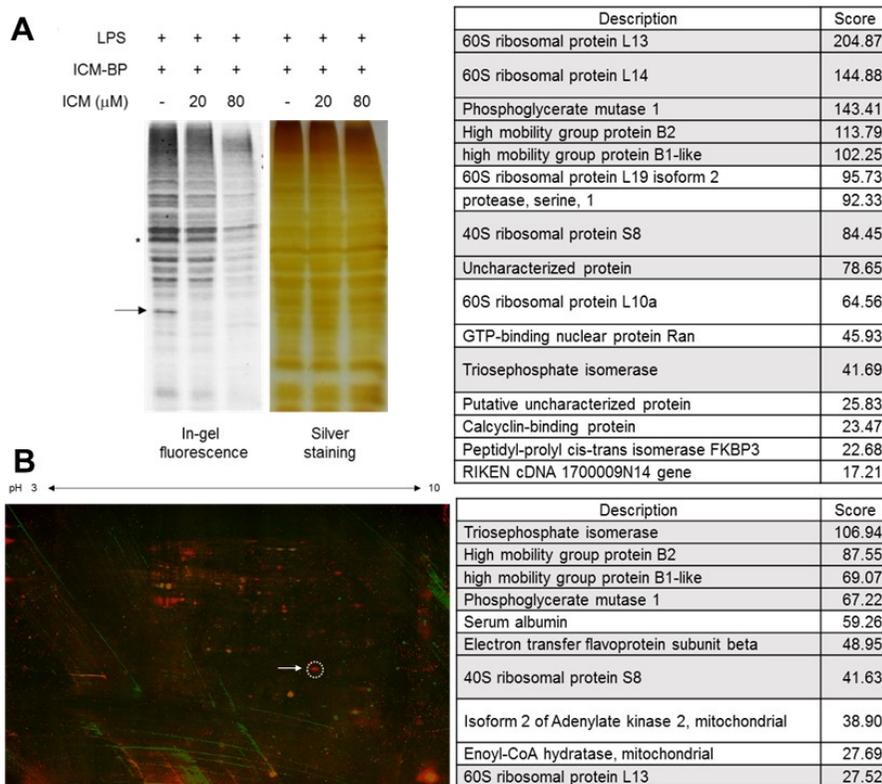


Figure 2.1.5. Target identification using ICM-BP by 1D and 2D gel analysis, and the mass analysis results. (A) Dose-dependent competitive labeling with ICM-BP (5 μ M) in live cells. (left) The band designated with arrow exhibited a significant difference even at a low concentration of the soluble competitor ICM. The list of proteins detected from gel bands in pointed arrow region by MS analysis (right). (B) 2D gel analysis for proteomes labeled with two different fluorophores in the presence of LPS stimulation. After incubation with compounds, cells were treated with LPS (100 ng/mL) for 30 min. The proteome photo-crosslinked with ICM-BP (20 μ M) was labeled with Cy5, and the proteomes treated with ICM-BP in the presence of ICM (20 μ M) was labeled with Cy3 to visualize non-specific binding events. Two proteomes labeled with Cy3 and Cy5 were mixed and subjected to 2D gel analysis. (left) The image represents the merged fluorescence between Cy5 (red) and Cy3 (green) as pseudo-color. The list of proteins detected from red spot gel (pointed white circle) by MS analysis (right). After cutting the indicated spot, it was subjected to trypsin digestion and analyzed by MS/MS. The proteins highlighted with gray background in the table indicate the proteins detected in both 1D and 2D analyses.

Next, we performed the target validation with a biophysical study and a loss-of-function study (Figure S2.1.3), also monitored the intracellular or extracellular translocation and post-transcriptional modification (PTM) of HMGB2, as well as isoform HMGB1 with 80% sequence homology,¹⁴ upon ICM treatment. (Figure S2.1.4) Modulation of PTM was simulated by computational docking study. The expected binding site of ICM is closed to the nuclear-localization signal (NLS) region of HMGBs, which have been known as a regulatory region of cellular translocation by PTM alternation. (Figure S2.1.5)

Considering the impact of microglia-mediated neurotoxicity in neuroinflammatory disease, we envisioned the therapeutic potential of ICM and

tested the effect of ICM on microglial activation *in vivo* using an LPS-induced mouse neuroinflammation model. Furthermore, to evaluate the protective role of ICM in the pathogenesis of neuroinflammatory disease, we utilized a murine experimental autoimmune encephalitis (EAE) model, which is an animal model of multiple sclerosis, a prototype organ-specific autoimmune and inflammatory disease in the CNS. We found that ICM administration for 15 days significantly reduced the progression of the disease, as determined by the EAE clinical score. (Figure S2.1.6)

Target Identification and Rational Optimization of Small-Molecule Enhancers for Glucose Uptake – Target ID Probe Design with the Least Modification of the Original Hit Compound¹⁵

Owing to the increasing morbidity associated with type 2 diabetes involving microvascular complications, the identification of effective methods to control hyperglycemia is considered a top priority in the biomedical community.¹⁶ Normal glucose levels can be maintained by regulating insulin sensitivity in skeletal muscles, glycogenolysis in the liver, and lipolysis in fat tissue.¹⁷ Therefore, the development of novel small molecules that regulate cellular glucose uptake or insulin sensitivity in muscles or fat tissues would provide promising candidates to treat type 2 diabetes and related complications. To facilitate the discovery of novel anti-diabetic agents, we envisioned the conjunction of target identification with phenotypic screening: 1) Image-based high-throughput screening (HTS)¹⁸ using a fluorescent glucose analogue, GB2, to identify small-molecule modulators of glucose uptake in myotubes;¹⁹ 2)

identification of the target protein of initial hits from a phenotypic assay by FITGE method;⁷ and 3) rational hit-to-lead optimization.

First, we performed the image-based HTS of a 3000-membered drug-like compound library derived from a privileged substructure-based diversity-oriented synthesis strategy²⁰ in a 96-well plate format. Fluorescence intensity of GB2 in the cytoplasm of the differentiated C2C12 myotubes was quantified from captured images in an automated fashion, which identified small-molecule enhancers of cellular glucose uptake. Among the initial hit compounds, we were particularly interested in two isoxazole-containing compounds (P29C06 and P29C07) because they selectively enhanced cellular glucose uptake in C2C12 myotubes, but not in undifferentiated C2C12 myoblasts. Their different glucose uptake capabilities related to the differentiation status of C2C12 cells were quantitatively investigated in myoblasts and myotubes by direct comparison with ampknone, a small-molecule glucose uptake enhancer that functions by the induction of AMP-activated protein kinase phosphorylation (Figure 2.1.7A, S2.1.7a and S2.1.7b).²¹ This phenotypic observation suggested that the cell differentiation-dependent enhancement of cellular glucose uptake may be caused by the elevated quantity of target proteins in myotubes compared to the pre-differentiated counterparts. Given the fact that both P29C06 and P29C07 also enhanced cellular glucose uptake in adipocytes, our initial hit compounds were considered to affect energy homeostasis.²² After identifying the initial hits from the phenotypic screening, we pursued the target identification of glucose uptake enhancers using FITGE technology, a fluorescence guided target identification method involving a photoaffinity

probe.⁷ Preliminary structure–activity relationship studies led us to modify P29C06 to generate a probe (P29C06-Az) for FITGE. As shown in Figure 2.1.7C, P29C06-Az contains aryl azide as a photoaffinity group and an acetylene moiety as a bioorthogonal handle for the visualization of cross-linked proteomes using a fluorescence reporter.²³ Despite structural changes, the probe retained its bioactivity in the myotubes and adipocytes (Figure S2.1.7c and S2.1.7d). To induce covalent tethering of P29C06-Az to its target proteins, the probe was incubated with live myotubes for 3 h, followed by UV irradiation for 30 min. The cells were then lysed and treated with an azide-containing Cy5 reporter to visualize P29C06-Az-protein complexes by click chemistry. To eliminate false-positive signals owing to non-specific labeling with photoaffinity probe, an excess of soluble competitor (P29C06) was added as a negative control in an identical procedure, where the Cy3 reporter was used for bioorthogonal labeling. Unfortunately, we were unable to identify target protein candidates in C2C12 myotubes, which might be due to the low expression levels of target proteins. On the other hand, we successfully identified a number of potential target proteins by mass analysis in differentiated 3T3L1 adipocytes (Figure 2.1.8, Tables S2.1.1 and S2.1.2). Among them, we identified PPAR γ as the target protein. In fact, PPAR γ is a member of the nuclear receptor family and has been shown to play important roles in energy homeostasis and glucose metabolism²⁴ associated with insulin sensitization and adipogenesis. Furthermore, its agonistic ligand, rosiglitazone, has been used as an oral therapeutic for the treatment of type 2 diabetes.²⁵

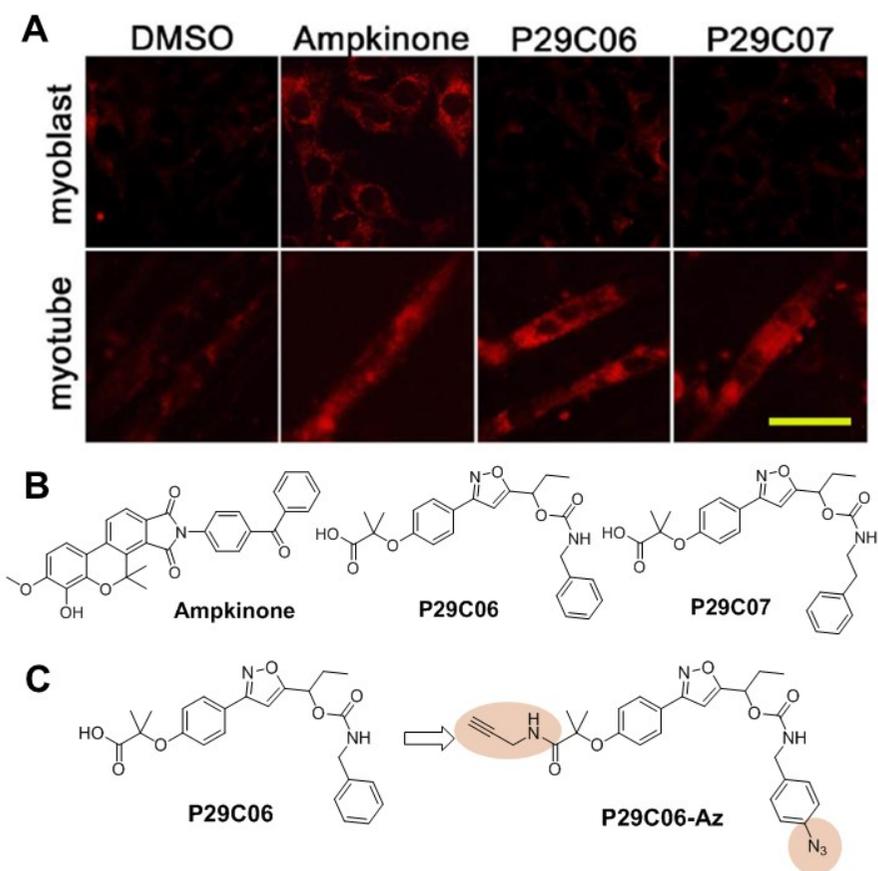


Figure 2.1.7. Target identification of small-molecule glucose uptake enhancers. (A) Fluorescence microscopic images of C2C12 myoblasts and myotubes 24 h after treatment with each compound using the GB2 procedure. Scale bar: 10 mm. (B) Chemical structures of ampkinone, P29C06, and P29C07. (C) Design of target ID probe P29C06-Az, with least structural modification of P29C06 to introduce both alkyne and aryl azide moieties.

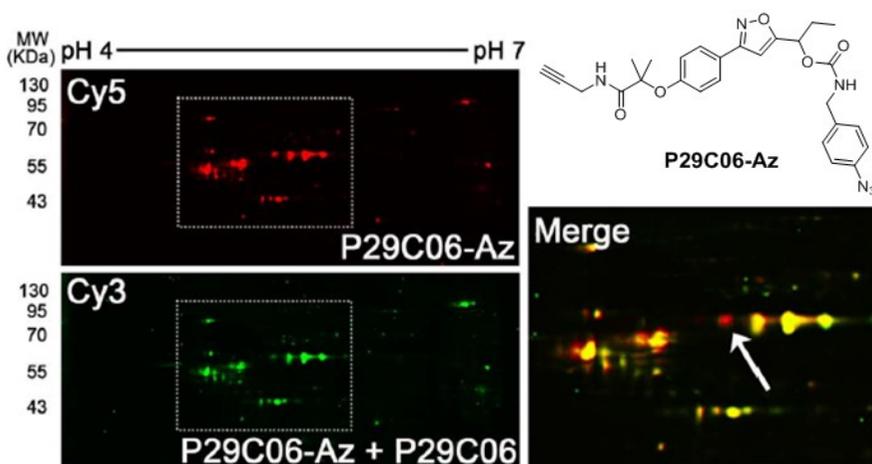


Figure 2.1.8. FITGE-based target identification of P29C06 using photoaffinity probe, P29C06-Az. Two-dimensional gel images were acquired in the Cy3 or Cy5 channel using a fluorescence scanner. The merged image shows the target protein selectively labeled in red (white arrow).

For the validation of the outcome of FITGE-based target identification, we subjected the initial hits to a cell-based PPAR γ -luciferase transactivation assay. As shown in Figure 2.1.9A, P29C06 and P29C07 exhibited partial agonism toward PPAR γ with moderate efficacy (EC₅₀ : 1.5 and 1.7 mM, respectively). The differential pattern of cellular glucose uptake in myotubes and myoblasts upon treatment with hit compounds might be due to the expression level of PPAR γ . Moderate activity of the initial hits drove us to pursue a rational drug discovery approach with the structural and functional information of PPAR γ . Based on the docking simulation, the initial hit compound P29C06 occupies the ligand-binding site²⁶ of PPAR γ with some extra space that can be utilized to enhance its efficacy as well as selectivity (Figure 2.1.9B). Therefore, we designed and synthesized a series of analogues to improve their efficacy in cellular glucose uptake (Table S2.1.3). First, we confirmed that the direct

analogues of initial hits, **1**{R₁,R₂}, having ethyl (E) group at the R₁ position and either benzyl (B) or phenethyl (PE) group at the R₂ position showed better transcriptional activity of PPAR γ . Next, we decorated **1**{E,B} and **1**{E,PE} with a second set of isocyanates to generate a 1,3-dicarbonyl moiety, which was predicted to be more tightly bound at the ligand-binding site of PPAR γ via an additional hydrogen bonding with Cys313 (Figure 2.1.9B and 2.1.9C). After further modification, we finally synthesized two enantiomers of the most active compound, **2**{E,B,35DMeOP}. Interestingly, each enantiomer exhibited different activity in the PPAR γ transactivation assay, with the R enantiomer (R35) being more potent than its S counterpart (S35; Figure 2.1.9D). Moreover, the EC₅₀ of R35 for PPAR γ transactivation was in the sub-nanomolar range, whereas that of S35 was 480-times lower (Figure 2.1.10A), which suggested that the binding of R35 favors the recruitment of co-activators with proper conformational changes of PPAR γ .²⁷

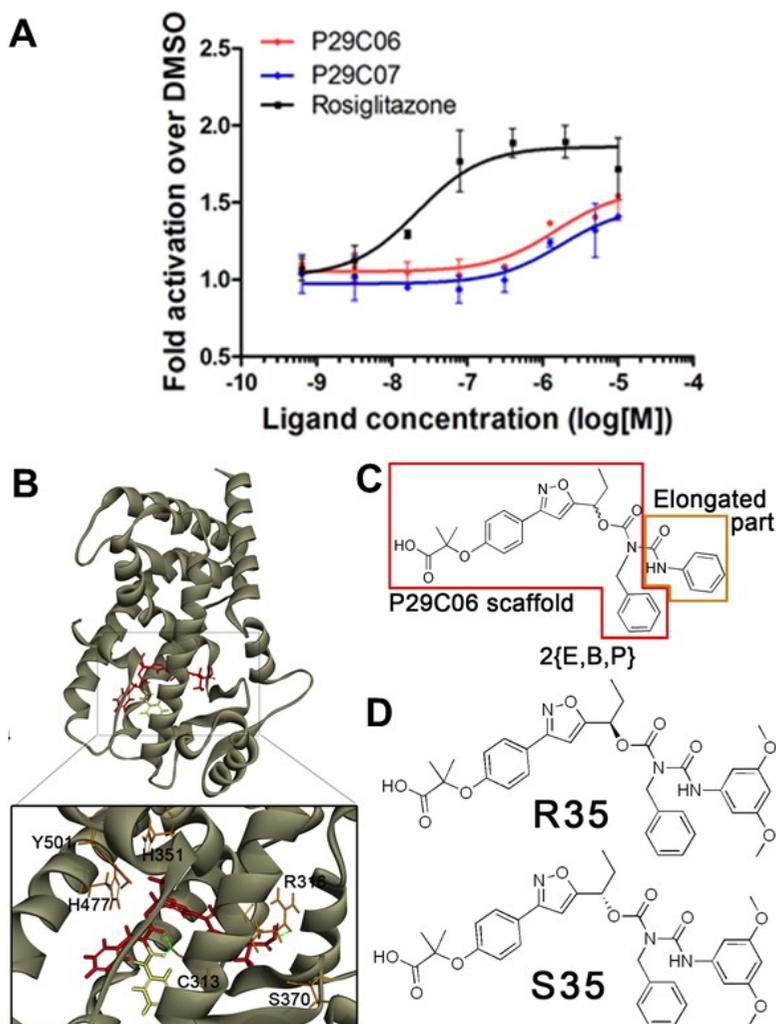


Figure 2.1.9. Target validation and structural optimization of primary hit compounds. (A) Transactivation profiles of rosiglitazone, P29C06, and P29C07 using a PPAR γ -derived reporter gene in 293T cells at 24 h after treatment ($n=3$). Error bars, s.d. (B) Docking analysis of 2{E,B,P} using crystal structure of PPAR γ ligand-binding domain (LBD) (PDB ID: 2hfp; Discovery Studio 1.7 [Accelrys] was used; Cys313 was intentionally tilted for better fit; hydrogen bonding is illustrated by the light green dashed line). (C) Design of optimized leads. Detailed synthetic procedure is described in Experimental section (2.1.5). (D) Chemical structures of R35 and S35, which were optimized PPAR γ agonists as well as PPAR γ phosphorylation inhibitors.

On the basis of two independent biophysical studies using surface plasmon resonance (SPR) spectroscopy and isothermal titration calorimetry (ITC), we confirmed that the binding affinity of R35 and S35 to the ligand binding domain (LBD) of PPAR γ was significantly enhanced from that of P29C06 (Figure 2.1.10B). These results indicated that the marginal transcriptional activity of P29C06 was due to its moderate binding affinity to the PPAR γ LBD. Along with this, R35 with an EC₅₀ of 0.38 nM showed a higher binding affinity to PPAR γ , compared to S35 and rosiglitazone. Collectively, we postulated the direct correlation of binding affinity with transactivation ability. However, we also observed that S35 has a comparable binding affinity to that of rosiglitazone, even though S35 is a much less potent agonist than rosiglitazone. Therefore, we suspected the mode of action of S35 to PPAR γ might be different from that of rosiglitazone. To test this hypothesis, we systematically compared our novel PPAR γ ligands, R35 and S35, for lead triage. First, both R35 and S35 increased glucose uptake to similar levels, which was measured using [¹⁴C]-2-deoxy-D-glucose in adipocytes (Figure 2.1.10C). We then sought to determine whether R35 and S35 might serve as inhibitors of the Cdk5-mediated phosphorylation of PPAR γ ,²⁸ which is an emerging mode of action for the development of novel antidiabetic agents without the side effects of rosiglitazone, a full PPAR γ agonist (Figure S2.1.8).^{25a,28,29}

Small-molecule ligands that inhibit the phosphorylation of PPAR γ at Ser273, such as MRL24, showed an insulin sensitizing effect in in vitro and in vivo systems by the regulation of gene sets related to the phosphorylation of PPAR γ .²⁹ To evaluate their biochemical function, we measured the inhibition

of Cdk5-mediated phosphorylation of PPAR γ at Ser273 upon treatment with various doses of either R35 or S35 using an in vitro enzymatic assay with purified Cdk5 and PPAR γ with MRL24 as a positive control.²⁸ As shown in Figure 2.1.10D, R35 effectively blocked Cdk5-mediated phosphorylation of PPAR γ with an IC₅₀ between 10 and 100 nm, while S35 showed a slightly lower activity than R35 (IC₅₀ between 100 and 1000 nm). We were confident that the phosphorylation of PPAR γ was specifically inhibited by ligand-mediated alteration of the interaction between Cdk5 and PPAR γ , rather than an inhibitory effect on general Cdk5 function.

In fact, the full agonism on PPAR γ triggers adipogenesis in fibroblasts,^{25b,d} which is a major adverse effect of rosiglitazone.^{28,29} As shown in Figure 2.1.10E, R35 potentiated adipocyte differentiation to a similar level of rosiglitazone, confirmed by the monitoring of cellular lipid accumulation with Oil Red O staining. However, in the case of S35, we observed extremely low levels of lipid accumulation, which was comparable to that of the vehicle control. These secondary biochemical evaluations imply that stereoisomeric differences in R35 and S35 may have different effects on gene expression, adipogenesis, and PPAR γ phosphorylation. R35 behaves as a conventional PPAR γ agonist while S35 deviated from agonism with high correlation in the phosphorylation inhibitory event unlike conventional PPAR γ agonists. Therefore, S35 may serve as a key structural clue for the development of new PPAR γ -related therapeutic agents to treat glucose homeostasis-related diseases without side effects.

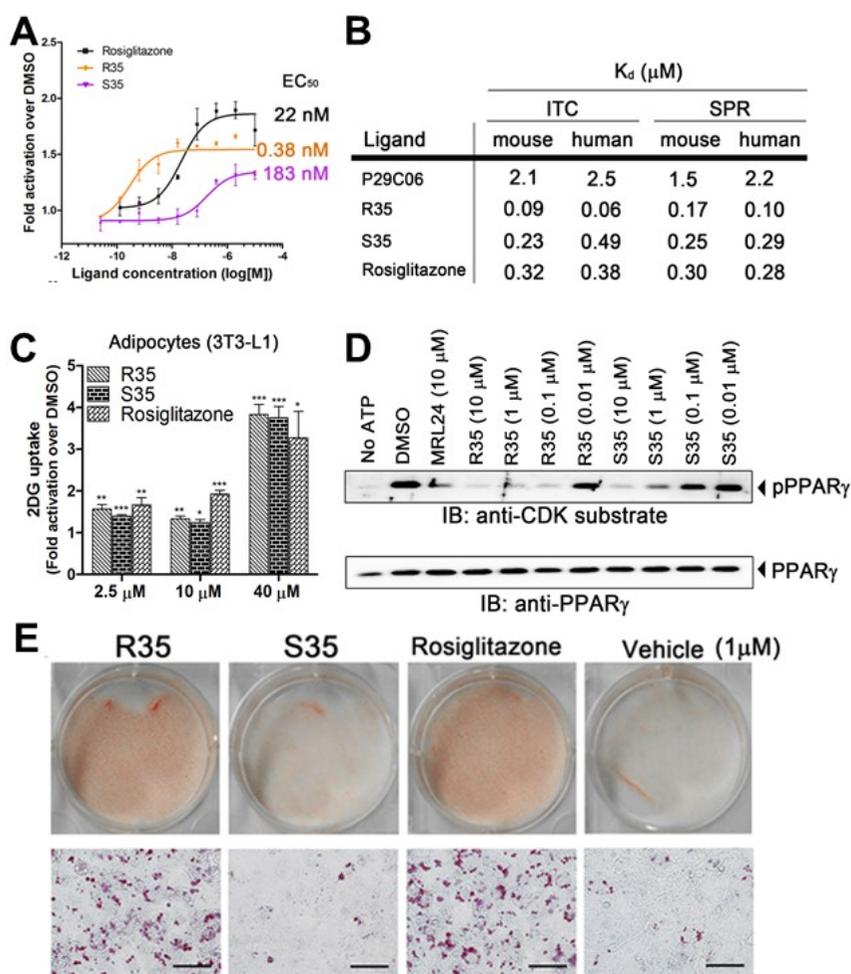


Figure 2.1.10. Identification of R35 and S35 as novel ligands of PPAR γ and differential phenotypes upon treatment with R35 and S35. (A) Transactivation profiles of rosiglitazone, R35, and S35 using a PPAR γ -derived reporter gene in 293T cells after 24 h treatment ($n=3$). Error bars, s.d. (B) Dissociation constants related to ligand-PPAR γ -LBD complex formation determined with ITC or SPR assay. (C) Glucose uptake profiles of fully differentiated 3T3L1 cells treated with rosiglitazone, R35, or S35 for 24 h, quantified using [^{14}C]-2-deoxy-D-glucose ($n=4$). (D) In vitro Cdk5 assay with MRL24, R35, and S35 using PPAR γ as a substrate. IB, immunoblot; NT, not treated; pPPAR γ , phosphorylated PPAR γ . (E) Differentiation status of 3T3L1 cells treated with rosiglitazone, R35, or S35 monitored by lipid accumulation with Oil Red O staining; scale bar, 10 mm.

2.1.4. Conclusion

In summary, target identification is the ‘crown jewel’ in chemical biology area by elucidating the biological mechanism of bioactive small molecule candidates. Herein, we described one of the successful target ID method, FITGE, and general strategy to design a probe for identifying target proteins. By looking over two studies we have been done, target ID was successfully done by rational construction of target ID probes based on the SAR of hit compounds and the least modification from the original candidate.

In the first target ID, the integration of phenotype-based screening with FITGE-based target identification led to the discovery of a new chemical entity, inflachromene (ICM). Target ID probes of ICM were rationally prepared based on the SAR of ICM, which enabled us to clarify its target protein, HMGB2 and another isoform HMGB1. We found that HMGB2 has a critical role in microglia-mediated neuroinflammation through the discovery of ICM and the subsequent identification of HMGB2 as the target of ICM. This convergent strategy and the subsequent mechanistic studies revealed that a novel small-molecule, a PTM-modulator of HMGBs, is an inhibitor of neuroinflammation with a broad window of therapeutic possibilities for various neuroinflammatory diseases.

In the second target ID case, we demonstrated the hit-to-lead optimization of new small-molecule enhancers of cellular glucose uptake, discovered from image-based phenotypic screening in myotubes. Our initial hit compounds were subjected to FITGE-guided target identification with simple structural modification to design the target ID probe, which revealed PPAR γ as the target

protein in adipocytes. The subsequent rational optimization of the initial hits generated a lead compound with high potency (4000-fold enhancement from initial hit). Secondary biophysical and biochemical studies revealed the stereoisomeric difference in our optimized PPAR γ ligands (R35 and S35) dictating their transcriptional activity, binding affinity, and inhibitory activity toward Cdk5-mediated phosphorylation of PPAR γ at Ser273. It turned out that S35 is a PPAR γ phosphorylation inhibitor with promising glucose uptake potential, while R35 is a highly potent conventional PPAR γ agonist; therefore, our PPAR γ ligands, especially S35, can be utilized for the development of new anti-diabetic agents without side effects. In light of high demands for a new class of therapeutic agents, the novel scaffolds identified from unbiased phenotypic screening and target identification can be a powerful resource for the development of first-in-class drugs.

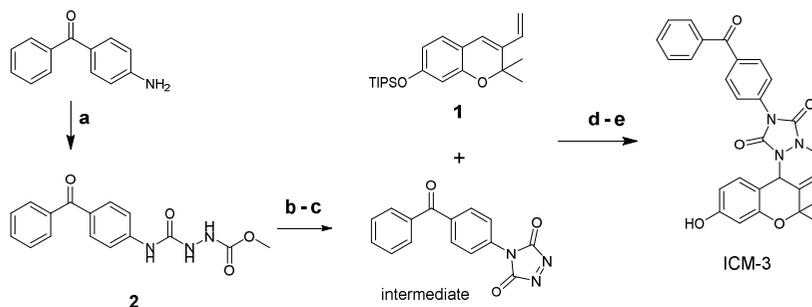
To conclude from these two successful target ID studies, we could establish a rational strategy to design target ID probes, and further FITGE-based target identification method efficiently enabled us to find out novel chemical entities for drug discovery.

2.1.5. Experimental Section

Detailed synthetic procedures of all new compounds (compound **2–4**, ICM-3, ICM-BP, **1**{R¹,R²}, **2**{R¹,R²,R³}, **R-5**{E}, and **P29C06-Az**) were described in the supporting informations of *Nat. Chem. Biol.* **2014**, *10*, 1055–1060 and

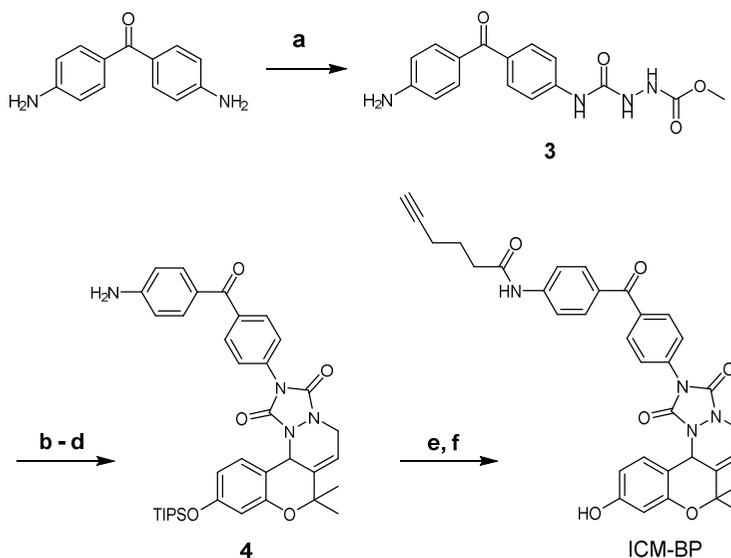
Angew. Chem. Int. Ed. **2014**, *53*, 5102–5106. The characterization of intermediates and other ICM analogues were previously reported.²¹

Supplementary Scheme 2.1.1. Synthesis of ICM-3.



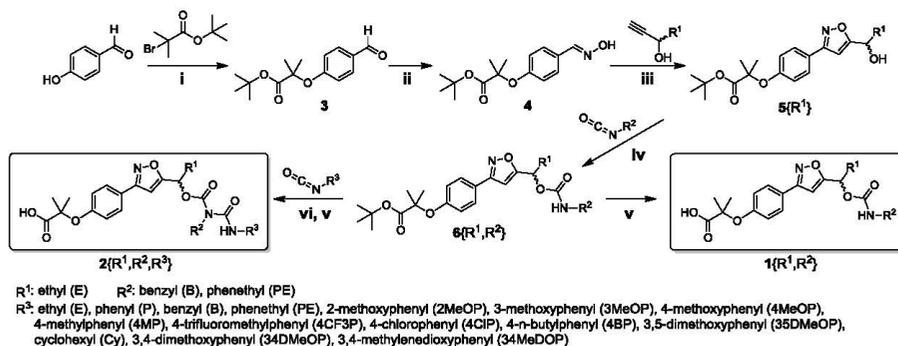
Reagents and conditions: (a) methyl carbamate, carbonyl diimidazole, THF/DMF (2:1), r.t. → 70 °C, 12 h; (b) K₂CO₃, MeOH, 70 °C, 12 h; (c) iodobenzene diacetate, THF, r.t., 1 h; (d) **1**, THF/toluene (1:1), r.t., 12 h; (e) HF/pyridine/THF (1:1:18), r.t., 6 h.

Supplementary Scheme 2.1.2. Synthesis of ICM-BP.



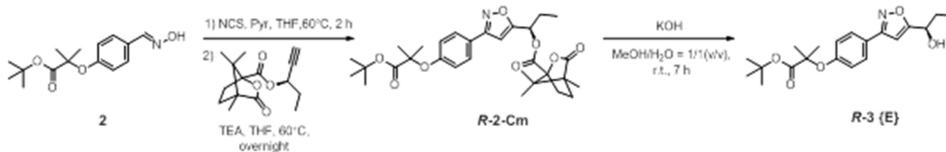
Reagents and conditions: (a) methyl carbazate, carbonyl diimidazole, THF/DMF (2:1), r.t. → 70 °C, 12 h; (b) K₂CO₃, MeOH, 70 °C, 12 h; (c) iodobenzene diacetate, THF, r.t., 1 h; (d) **1**, THF/toluene (1:1), r.t., 12 h; (e) 5-hexynoic acid, oxalyl chloride, pyridine, 0 °C, 2 h; (f) HF/pyridine/THF(1:1:18), r.t., 6 h.

Supplementary Scheme 2.1.3. Synthesis of **1**{R¹,R²} and **2**{R¹,R²,R³}.

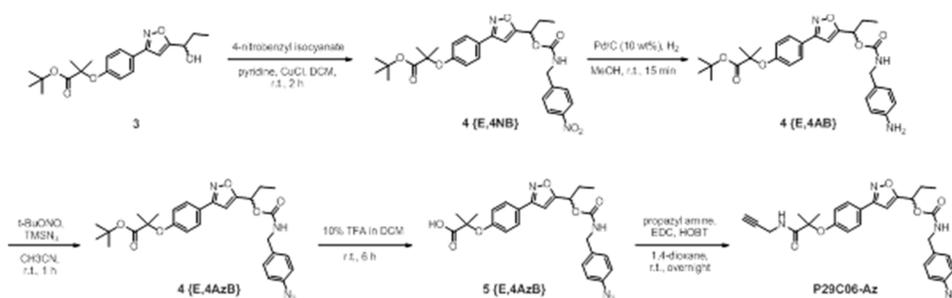


Reagents and conditions: (i) K₂CO₃, MgSO₄, DMF, 100 °C, 22 h, 99%; (ii) hydroxylamine, Na₂CO₃, aq. ethanol, 65 °C, 12 h, 85%; (iii) NCS, pyridine, followed by acetylene, TEA, THF, 60 °C, 2 h, 62%; (iv) pyridine, CuCl, DCM, r.t., 2 h, 90–100%; (v) TFA, DCM, r.t., 6 h, 80–97%; (vi) TEA, CuCl, DCM, r.t. 12 h, 50–85%.

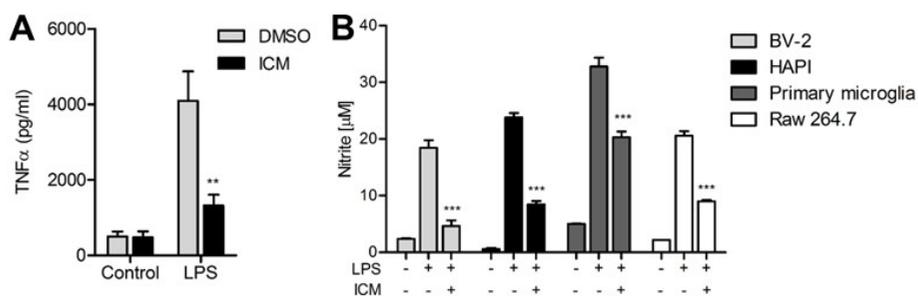
Supplementary Scheme 2.1.4. Synthetic procedure for **R-5**{E}.



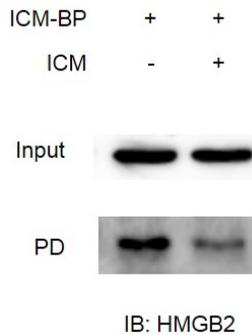
Supplementary Scheme 2.1.5. Synthetic procedure of P29C06-Az.



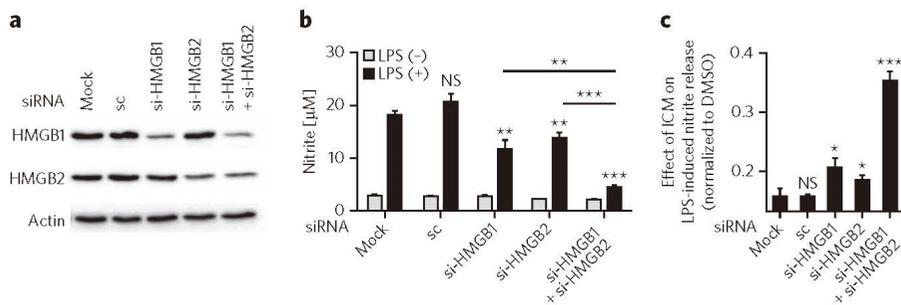
2.1.6. Supporting Information



Supplementary Figure 2.1.1. (A) LPS-induced TNF α secretion in BV-2 cells was measured by ELISA. (B) The effects of ICM on various cell lines, including BV-2 (mouse microglia), HAPI (rat microglia), primary microglia, and RAW 264.7 (mouse macrophages). Data represent the mean and s.d. (** $P < 0.001$; *** $P < 0.0001$; ns, no significant compared to DMSO).

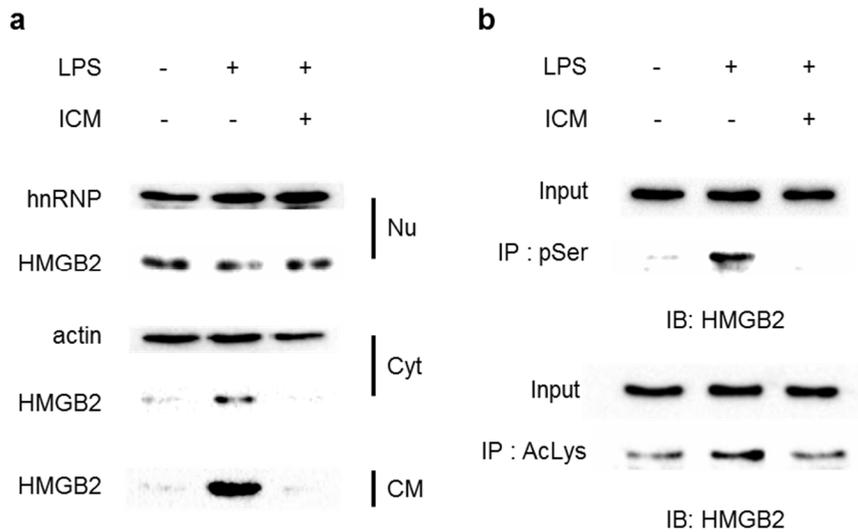


Supplementary Figure 2.1.2. Pull-down (PD) assay was performed with biotin-labeled proteome and streptavidin beads. ICM-BP (5 μ M) with or without ICM competition (20 μ M) was subjected to PD and immunoblotting (IB) for HMGB2.

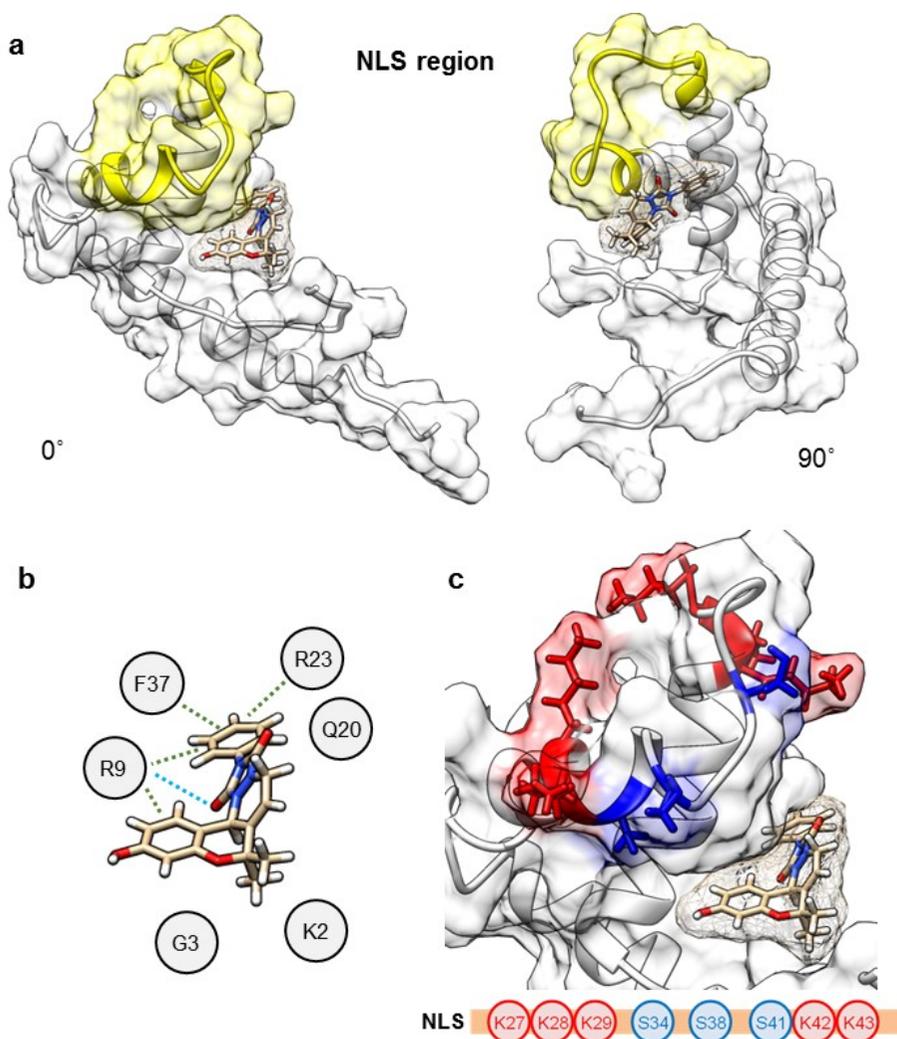


Supplementary Figure 2.1.3. Functional validation of HMGB2 as a direct binding target of ICM.

(a) Representative western blot results in triplicate confirming siRNA-mediated knockdown of HMGB1 and HMGB2 in BV-2 cells (mock, control transfection; sc, scrambled siRNA; si-HMGB1, siRNA for HMGB1; si-HMGB2, siRNA for HMGB2). (b) The loss-of-function effect by siRNA knockdown with either si-HMGB1 or si-HMGB2 and double knockdown with both si-HMGB1 and si-HMGB2 confirmed by LPS-induced nitrite release in BV-2 cells. (c) The HMGB-dependent effect of ICM. LPS-induced nitrite release after siRNA transfection. ICM (5 μ M) treatment was normalized to the individual DMSO treatment. All data represent the mean and s.d. in triplicate. * $P < 0.05$, ** $P < 0.001$, *** $P < 0.0001$ compared to mock transfection, as calculated by Student's *t*-test NS, not significant compared to mock transfection. Asterisks above the lines indicate significance between single and dual knockdown.

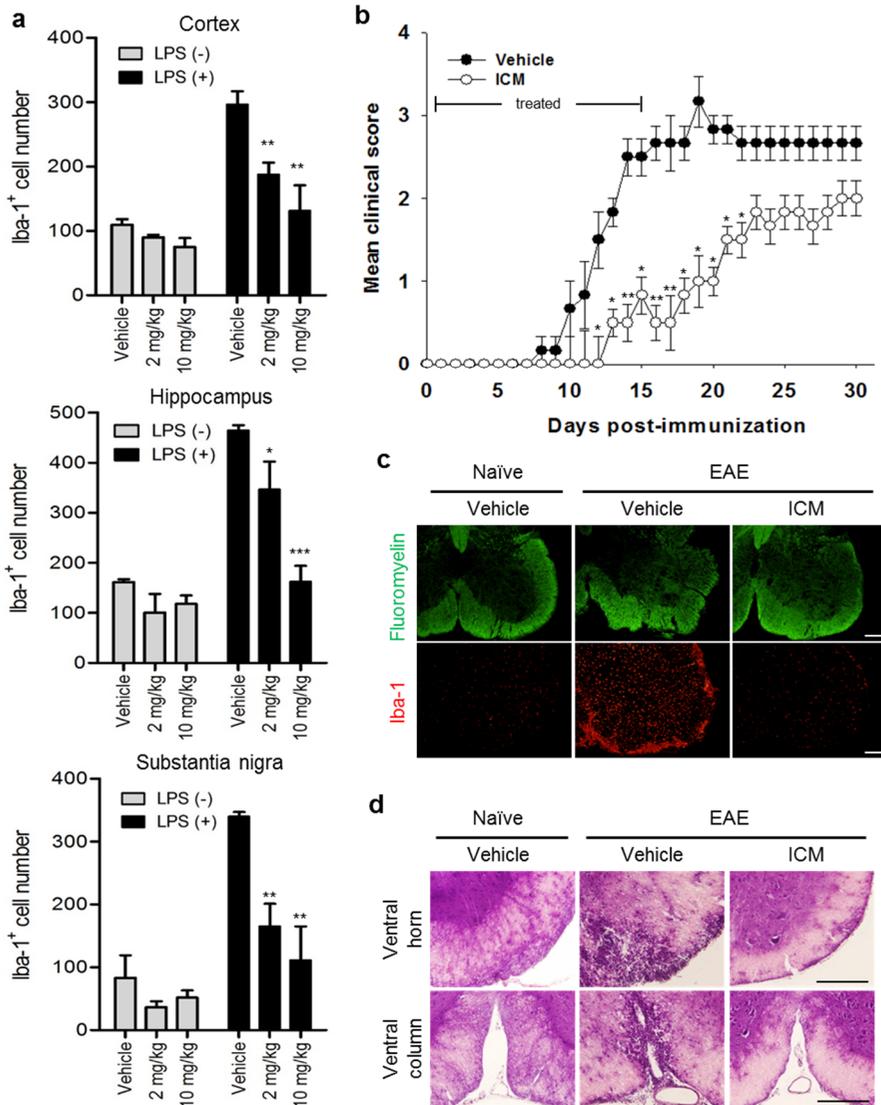


Supplementary Figure 2.1.4. Mode-of-action study of ICM for perturbing the inflammatory function of HMGB2. (a) The effect of ICM on the translocation of HMGB2 from the nucleus (Nu) to the cytoplasm (Cyt) and extracellular milieu in conditioned media (CM). (b) Anti-phosphoserine (pSer) and acetylated lysine (AcLys) immunoprecipitation (IP) followed by immunoblotting (IB) for HMGB2 indicated the effect of ICM on the post-translational modification of HMGB2. BV-2 cells were treated with ICM (10 μ M) prior to LPS (200 ng/mL), and analyzed by subcellular fractionation or IP assay. hnRNP and actin were detected as housekeeping proteins in nucleus and cytoplasm, respectively.



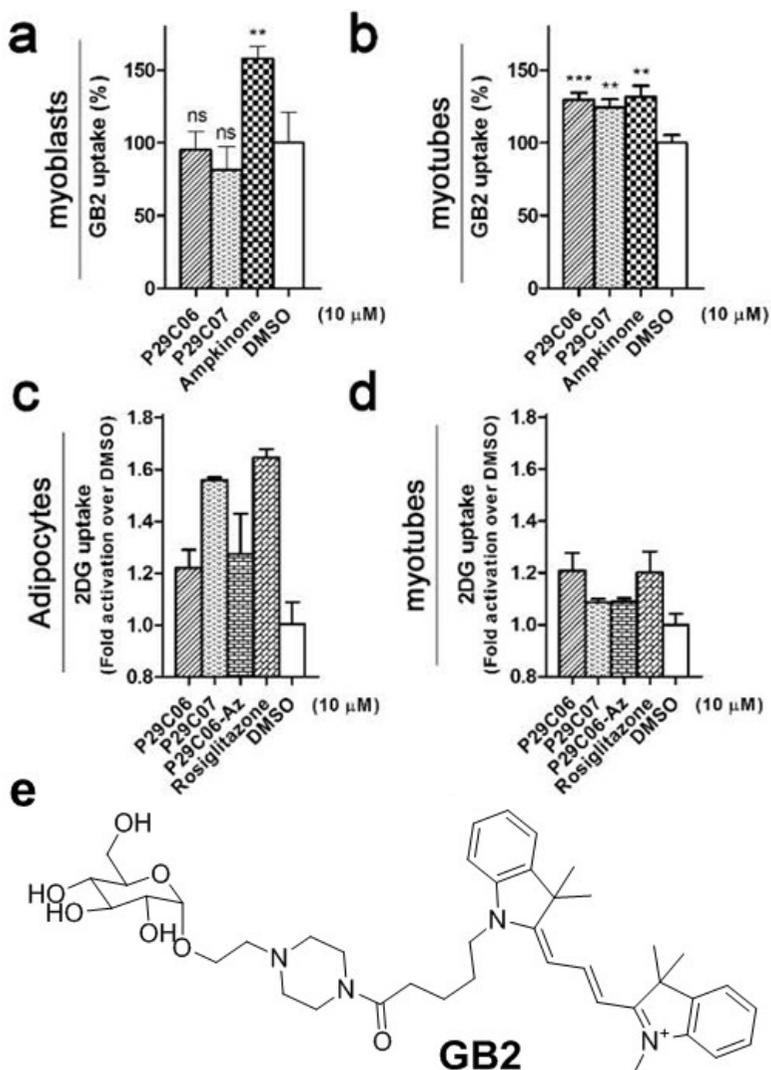
Supplementary Figure 2.1.5. Prediction of the relationship between ICM-binding event with HMGB1 and PTM at the NLS region of HMGB1. (a) Spatial proximity between ICM and NLS region of HMGB1 in the docking simulation. The NLS region of HMGB1 was indicated as yellow color. (b) Interaction of ICM with adjacent amino acids at the cavity of HMGB1. Amino acid residues within the range of 4 Å from ICM were illustrated around ICM. Individual interaction between ICM and amino acid residues was described by dotted line; pi-pi interaction with F37 as a green, cation-pi interactions with R9 and R23 as a green, and hydrogen bonding interaction with R9 as a skyblue. (c) Expanded image of the binding site of HMGB1 with ICM

(up) and the NLS region of HMGB1 highlighted with Lys and Ser (bottom). As a potential site for acetylation and phosphorylation, several Lys and Ser residues are located at the NLS region (sequence 27~43) in box A domain of HMGB1. Lys and Ser residues are indicated by red and blue, respectively.



Supplementary Figure 2.1.6. The therapeutic effect of ICM *in vivo*. (a) The effect of ICM on

microglial activation in mouse brain after LPS injection (i.p.). Quantitative analysis of Iba-1 (the marker of microglial activation) immunofluorescence staining in cortex, hippocampus, and substantia nigra (SN) regions. Data represent the mean and s.d. ($n = 3$, * $P < 0.01$, ** $P < 0.001$, *** $P < 0.0001$ compared to the vehicle). (b) Mean clinical scores from the EAE induction in C57BL/6 mice. The animals were administered daily with vehicle or ICM (10 mg/kg) for 15 days after immunization with the MOG₃₅₋₅₅ peptide and pertussis toxin. The immunized mice were monitored and scored daily, and the severity of EAE was determined in a blinded fashion. Data represent the mean and s.e.m. ($n = 6$, * $P < 0.05$, ** $P < 0.01$ compared to the vehicle). (c) Histological analysis of ICM effects on the EAE model. Lumbar spinal cords from each group were removed at the disease peak time (day 15 of EAE induction). Frozen sections of lumbar spinal cords were stained with fluoromyelin for myelin (*upper*) and anti-Iba-1 antibody for microglial activation (*lower*), compared with pre-immunized mice (Naïve). (d) Frozen sections in c were subjected to H&E staining to show inflammatory lesions. Scale bar is 200 μm .



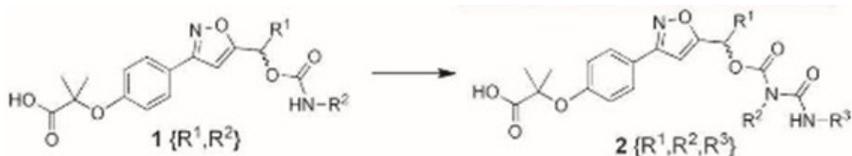
Supplementary Figure S2.1.7. Glucose uptake profiles of myoblasts (a) or myotubes (b) upon treatment with **P29C06**, **P29C07**, or ampkinone for 24 h quantified using GB2 as a fluorescence glucose uptake tracer; ** $P < 0.001$, *** $P < 0.0001$; ns, not significant. (c, d) Glucose uptake profile in fully differentiated adipocytes (3T3-L1 cells) or myotubes (C2C12 cells) upon treatment with rosiglitazone, **P29C06**, **P29C07**, or **P29C06-Az** for 24 h quantified by [14 C]-2-deoxy-D-glucose. Error bars, s.e.m. (e) Chemical structure of GB2.^{19a}

N	Accession	Description	S	C (%)	Pro	UP	P	PSMs	AA	MW (kDa)	calc. pI
1	259090165	Chain A, Structural Basis Of Pp2a And Sgo Interaction	55.95	26.53	4	9	13	18	588	65.1	5.11
2	148690914	nucleobindin 1, isoform CRA_b [Mus musculus]	46.26	32.09	6	13	14	16	455	53.0	5.07
3	148675845	RIKEN cDNA 1810073G14, isoform CRA_b [Mus musculus]	41.07	35.71	6	9	11	14	490	53.5	4.86
4	2078001	vimentin [Mus musculus]	39.23	29.82	25	10	12	14	446	51.5	5.02
5	33440473	Protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform [Mus musculus]	31.24	17.80	6	5	9	11	601	65.9	5.07
6	148698219	karyopherin (importin) alpha 6 [Mus musculus]	22.77	17.45	7	7	8	8	533	59.6	4.98
7	76779273	Hspd1 protein, partial [Mus musculus]	21.37	17.12	8	7	7	7	555	59.4	8.07
8	148683551	karyopherin (importin) alpha 4 [Mus musculus]	18.16	14.20	2	3	4	5	521	57.9	4.94
9	148695990	mCG6540 [Mus musculus]	14.90	14.68	6	4	4	5	470	52.0	5.21
10	148670616	RIKEN cDNA 1110036O03 [Mus musculus]	14.21	12.87	2	5	5	5	443	51.1	4.81
11	81896299	RecName: Full=Protein in prune homolog; AltName: Full=PRUNEM1	13.93	14.10	3	4	4	5	454	50.2	5.11
12	148704153	karyopherin (importin) alpha 3 [Mus musculus]	8.40	7.63	3	2	3	3	498	55.1	4.83
13	148686957	mCG122657 [Mus musculus]	7.15	7.73	1	2	2	2	440	49.2	6.21
14	148699710	paralemmin, isoform CRA_b [Mus musculus]	6.83	13.27	7	2	3	3	339	36.7	4.84
15	148676670	heat shock 70kD protein 5 (glucose-regulated protein), isoform CRA_b [Mus musculus]	5.44	6.11	12	1	2	2	507	56.3	5.14
16	229577163	epidermal growth factor receptor substrate 15 isoform B [Mus musculus]	5.29	4.12	5	2	2	2	583	64.3	4.40
17	89574015	mitochondrial ATP synthase, H+ transporting F1 complex beta subunit [Mus musculus]	5.22	7.19	13	2	2	2	445	48.0	5.02
18	148701192	epidermal growth factor-containing fibulin-like extracellular matrix protein 2, isoform CRA_a [Mus musculus]	5.19	4.51	3	1	2	2	443	49.4	4.89
19	148677882	dyactin 4, isoform CRA_d [Mus musculus]	4.77	8.23	6	1	2	2	498	56.2	7.62
20	38328414	Ppp4r2 protein, partial [Mus musculus]	4.55	7.69	4	1	2	2	390	43.2	4.53
21	187960105	peroxisome proliferator-activated receptor gamma isoform 1 [Mus musculus]	4.21	4.63	9	1	2	2	475	54.4	6.68
22	191765	alpha-fetoprotein, partial [Mus musculus]	2.74	3.11	5	1	1	1	418	47.2	5.66
23	1405933	M2-type pyruvate kinase [Mus musculus]	2.67	2.64	7	1	1	1	531	57.9	7.47
24	148674283	ribophorin II [Mus musculus]	2.60	3.87	6	1	1	1	594	64.8	5.55
25	47847474	mFJ00216 protein [Mus musculus]	2.35	4.31	4	1	1	1	418	46.1	5.50

Supplementary Table S2.1.1. List of identified proteins from FITGE.

Sequence	PSMs	Pro	Protein Group Accessions	ΔC_n	XCorr	Probability	Charge	MH+ [Da]	ΔM [ppm]	RT [min]
SVEAVQETETAYK	1	9	187960105	0.0000	2.12	0.00	2	1466.74004	1.82	26.94
ADPMVADYK	1	8	187960105	0.0000	2.09	0.00	2	1009.46825	2.27	19.18

Supplementary Table S2.1.2. Information of the identified peptides of PPAR γ .



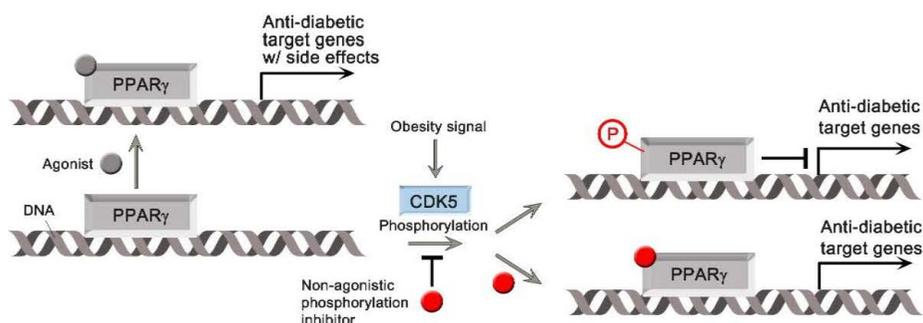
R¹: ethyl (E) R²: benzyl (B), phenethyl (PE)

R³: ethyl (E), phenyl (P), benzyl (B), phenethyl (PE), 2-methoxyphenyl (2MeOP), 3-methoxyphenyl (3MeOP), 4-methoxyphenyl (4MeOP), 4-methylphenyl (4MP), 4-trifluoromethylphenyl (4CF3P), 4-chlorophenyl (4ClP), 4-n-butylphenyl (4BP), 3,5-dimethoxyphenyl (35DMeOP), cyclohexyl (Cy), 3,4-dimethoxyphenyl (34MeOP), 3,4-methylenedioxyphenyl (34MeDOP)

Entry	Compounds	PPAR α	PPAR δ	PPAR γ
1	1 {E, B}	> 20000	IA ^[c]	1510
2	1 {E, PE}	IA	IA	1730
3	2 {E, B, E}	IA	IA	201
4	2 {E, B, P}	IA	IA	6.01
5	2 {E, B, B}	364	IA	51.6
6	2 {E, B, PE}	3830 ^[b]	> 10000	42.2
7	2 {E, PE, E}	IA	479 ^[b]	> 10000
8	2 {E, PE, P}	IA	IA	180
9	2 {E, PE, B}	IA	IA	> 10000
10	2 {E, PE, PE}	7160 ^[b]	383 ^[b]	382
11	2 {E, B, 2MeOP}	IA	17.2 ^[b]	421
12	2 {E, B, 3MeOP}	IA	IA	3.9
13	2 {E, B, 4MeOP}	4290 ^[b]	IA	12.6
14	2 {E, B, 4MP}	IA	IA	9.2
15	2 {E, B, 4CF3P}	1710 ^[b]	IA	41.7
16	2 {E, B, 4ClP}	IA	IA	78.7
17	2 {E, B, 4BP}	266 ^[b]	IA	7.8
18	2 {E, B, 34DMeOP}	IA	IA	23.8
19	2 {E, B, 35DMeOP}	IA	IA	3.3
20	2 {E, B, 34MeDOP}	IA	IA	3.8
21	WY14643	669	IA	> 10000
22	GW501516	> 10000	1.4	> 10000
23	Rosiglitazone	> 10000	> 10000	22

[a] units: nM. [b] IC₅₀. [c] inactive.

Supplementary Table S2.1.3. Transactivation profiles of WY14643, GW501516, rosiglitazone, and compounds 1 {R¹, R²} and 2 {R¹, R², R³} using a PPAR-derived reporter gene in 293T cells at 24 h after treatment with each compound.^[a]



Supplementary Figure S2.1.8. Regulation of PPAR γ activity with a conventional PPAR γ agonist or small-molecule inhibitor of Cdk5-mediated phosphorylation of S273 of PPAR γ . The inhibition of Cdk5-mediated phosphorylation of PPAR γ is considered to be highly beneficial for the treatment of type II diabetes without side effects via selective boosting of PPAR γ transcriptional activity.

2.1.7. References and Notes

- † Portions of this chapter have been previously reported, see: (a) Lee, S.; Nam, Y.; Koo, J. Y.; Lim, D.; Park, J.; Ock, J.; Suk, K.; Park, S.B. *Nat. Chem. Biol.* **2014**, *10*, 1055–1060. (b) Koh, M.; Park, J.; Koo, J. Y.; Lim, D.; Cha, M. Y.; Jo, A.; Choi, J. H.; Park, S. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 5102–5106.
- (1) (a) Czechtizky, W.; Hamley, P. *Small Molecule Medicinal Chemistry: Strategies and Technologies*, 1st ed.; John Wiley & Sons, Inc., 2016.
 (b) Eder, J.; Sedrani, R.; Wiesmann, C. *Nat. Rev. Drug Discov.* **2014**, *13*, 577–587.
- (2) Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. *Angew. Chem., Int. Ed.* **2013**, *52*, 2744–2792.

- (3) (a) Pan, S.; Zhang, H.; Wang, C.; Yao, S. C.; Yao, S. Q. *Nat. Prod. Rep.* **2016**, *33*, 612–620. (b) Moffat, J. G.; Rudolph, J.; Bailey, D. *Nat. Rev. Drug Discov.* **2014**, *13*, 588–602.
- (4) Schenone, M.; Dancik, V.; Wagner, B. K.; Clemons, P. A. *Nat. Chem. Biol.* **2013**, *9*, 232–240.
- (5) Simon, G. M.; Niphakis, M. J.; Cravatt, B. F. *Nat. Chem. Biol.* **2013**, *9*, 200–205.
- (6) Ong, S.-E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D.B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376–386.
- (7) Park, J.; Oh, S.; Park, S.B. *Angew. Chem. Int. Ed.* **2012**, *51*, 5447–5451.
- (8) (a) Bi, X.; Schmitz, A.; Hayallah, A. M.; Song, J. N.; Famulok, M. *Angew. Chem., Int. Ed.* **2008**, *47*, 9565–9568. (b) Liu, K.; Shi, H.; Xiao, H.; Chong, A. G. L.; Bi, X.; Chang, Y. T.; Tan, K. S. W.; Yada, R. Y.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2009**, *48*, 8293–8297. (c) Ban, H. S.; Shimizu, K.; Minegishi, H.; Nakamura, H.; *J. Am. Chem. Soc.* **2010**, *132*, 11870–11871; (d) Eirich, J.; Orth, R.; Sieber, S. A. *J. Am. Chem. Soc.* **2011**, *133*, 12144–12153. (e) Johnson, K.; Zhu, S.; Tremblay, M. S.; Payette, J. N.; Wang, J.; Bouchez, L. C.; Meeusen, S.; Althage, A.; Cho, C. Y.; Wu, X.; Schultz, P. G. *Science*, **2012**, *336*, 717–721. (f) MacKinnon, A. L.; Garrison, J. L.; Hegde, R. S.; Taunton, J. *J. Am. Chem. Soc.* **2007**, *129*, 14560–14561. (g) Shi, H.; Zhang, C.-J.; Chen, G. Y. J.; Yao, S. Q. *J. Am. Chem. Soc.* **2012**, *134*, 3001–3014. (h) Cisar, J. S.; Cravatt, B. F. *J. Am. Chem. Soc.* **2012**, *134*, 10385–

10388.

- (9) Lee, S.; Nam, Y.; Koo, J.Y.; Lim, D.; Park, J.; Ock, J.; Suk, K.; Park, S.B. *Nat. Chem. Biol.* **2014**, *10*, 1055–1060.
- (10) Block, M. L.; Zecca, L.; Hong, J-S. *Nat. Rev. Neurosci.* **2007**, *8*, 57–69.
- (11) Oh, S.; Park, S.B. *Chem. Commun.* **2011**, *47*, 12754–12761.
- (12) Lotze, M. T.; Tracey, K. J. *Nat. Rev. Immunol.* **2005**, *5*, 331–342.
- (13) Sims, G. P.; Rowe, D. C.; Rietdijk, S. T.; Herbst, R.; Coyle, A. J. *Annu. Rev. Immunol.* **2010**, *28*, 367–388.
- (14) Yang, H.; Tracey, K. J. *Biochim. Biophys. Acta* **2010**, *1799*, 149–156.
- (15) Koh, M.; Park, J.; Koo, J.Y.; Lim, D.; Cha, M.Y.; Jo, A.; Choi, J.H.; Park, S. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 5102–5106.
- (16) (a) Cohen, E.; Dillin, A. *Nat. Rev. Neurosci.* **2008**, *9*, 759–767. (b) Nathan, D. M.; Buse, J. B.; Davidson, M. B.; Ferrannini, E.; Holman, R. R.; Sherwin, R.; Zinman, B. *Diabetes Care* **2009**, *32*, 193–203. (c) Semenkovich, C. F. *J. Clin. Invest.* **2006**, *116*, 1813–1822.
- (17) Kars, M.; Yang, L.; Gregor, M. F.; Mohammed, B. S.; Pietka, T. A.; Finck, B. N.; Patterson, B. W.; Horton, J. D.; Mittendorfer, B.; Hotamisligil, G. S.; Klein, S. *Diabetes* **2010**, *59*, 1899–1905.
- (18) Jo, A.; Park, J.; Park, S. B. *Chem. Commun.* **2013**, *49*, 5138–5140.
- (19) (a) Lee, H. Y.; Lee, J. J.; Park, J.; Park, S. B. *Chem. Eur. J.* **2011**, *17*, 143–150. (b) Park, J.; Lee, H. Y.; Cho, M. H.; Park, S. B. *Angew. Chem. Int. Ed.* **2007**, *46*, 2018–2022. (c) Tian, Y. S.; Lee, H. Y.; Lim, C. S.; Park, J.; Kim, H. M.; Shin, Y. N.; Kim, E. S.; Jeon, H. J.; Park,

- S. B.; Cho, B. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 8027–8031.
- (20) Oh, S.; Park, S. B. *Chem. Commun.* **2011**, *47*, 12754–12761.
- (21) Oh, S.; Kim, S. J.; Hwang, J. H.; Lee, H. Y.; Ryu, M. J.; Park, J.; Kim, S. J.; Jo, Y. S.; Kim, Y. K.; Lee, C. H.; Kweon, K. R.; Shong, M.; Park, S. B. *J. Med. Chem.* **2010**, *53*, 7405–7413.
- (22) Rajala, M. W.; Scherer, P. E. *Endocrinology* **2003**, *144*, 3765–3773.
- (23) Staros, J. V. *Trends Biochem. Sci.* **1980**, *5*, 320–322.
- (24) Ryan, K. K.; Li, B.; Grayson, B. E.; Matter, E. K.; Woods, S. C.; Seeley, R. J. *Nat. Med.* **2011**, *17*, 623–626.
- (25) (a) Ahmadian, M.; Suh, J. M.; Hah, N.; Liddle, C.; Atkins, A. R.; Downes, M.; Evans, R. M. *Nat. Med.* **2013**, *19*, 557–566. (b) Albrektsen, T.; Frederiksen, K. S.; Holmes, W. E.; Boel, E.; Taylor, K.; Fleckner, J. *Diabetes* **2002**, *51*, 1042–1051. (c) Semple, R. K.; Chatterjee, V. K.; Rahilly, S. O. *J. Clin. Invest.* **2006**, *116*, 581–589. (d) Tontonoz, P.; Hu, E.; Spiegelman, B. M. *Cell* **1994**, *79*, 1147–1156.
- (26) Nolte, R. T.; Wisely, G. B.; Westin, S.; Cobb, J. E.; Lambert, M. H.; Kurokawa, R.; Rosenfeld, M. G.; Willson, T. M.; Glass, C. K.; Milburn, M. V. *Nature* **1998**, *395*, 137–143.
- (27) Kallenberger, B. C.; Love, J. D.; Chatterjee, V. K.; Schwabe, J. W. *Nat. Struct. Biol.* **2003**, *10*, 136–140.
- (28) Choi, J. H.; Banks, A. S.; Estall, J. L.; Kajimura, S.; Bostrom, P.; Laznik, D.; Ruas, J. L.; Chalmers, M. J.; Kamenecka, T. M.; Bluher, M.; Griffin, P. R.; Spiegelman, B. M. *Nature* **2010**, *466*, 451–456.
- (29) Choi, J. H.; Banks, A. S.; Kamenecka, T. M.; Busby, S. A.; Chalmers,

M. J.; Kumar, N.; Kuruvilla, D. S.; Shin, Y.; He, Y.; Bruning, J. B.;
Marciano, D. P.; Cameron, M. D.; Laznik, D.; Jurczak, M. J.; Schurer,
S. C.; Vidovic', D.; Shulman, G. I.; Spiegelman, B. M.; Griffin, P. R.
Nature **2011**, *477*, 477–481.

- (30) GraCa, X.; De Luca, A.; Sang, N.; Fu, Y.; Claudio, P. P.; Rosenblatt,
J.; Morgan, D. O.; Giordano, A. *Proc. Natl. Acad. Sci. USA* **1994**, *91*,
3834–3838.

2.2. Investigation of Specific Binding Proteins to Photoaffinity Linkers for Efficient Deconvolution of Target Protein

2.2.1. Abstract

Photoaffinity-based target identification has received recent attention as an efficient research tool for chemical biology and drug discovery. The major obstacle of photoaffinity-based target identification is the nonspecific interaction between target identification probes and non-target proteins. Consequently, the rational design of photoaffinity linker has been spotlighted for successful target identification. These nonspecific interactions have been considered as random events and therefore no systematic investigation has been conducted regarding nonspecific interactions between proteins and photoaffinity linkers. Herein, we report the protein-labeling analysis of photoaffinity linkers containing three photoactivatable moieties: benzophenone, diazirine, and aryl azide. Each photoaffinity linker binds to a different set of proteins in a structure-dependent manner, in contrast to the previous conception. The list of proteins labeled by each photoaffinity linker was successfully used to eliminate the nonspecific binding proteins from target candidates, thereby increasing the success rate of target identification.

2.2.2. Introduction

In the post-genomic era, phenotype-based drug discovery has emerged as a promising approach for the discovery of first-in-class therapeutic agents.¹⁻⁴ Even though this emerging technique enables innovation in the discovery of bioactive small molecules with novel mechanisms of action, phenotype-based approaches require successful target identification (ID) and deconvolution,⁵ which is the rate-limiting process in drug discovery.⁶ Various studies have tried to develop new strategies for the identification of target proteins of bioactive small molecules.⁷⁻¹⁵ The conventional method for target ID employs affinity-based pull-down assays by exploiting non-covalent interactions between bioactive small molecules and target proteins in cell lysates. But the proteins identified using this method are severely affected by experimental conditions such as buffers, washing conditions, incubation times, and temperature.¹⁶ Nonspecific protein binding to the affinity bead itself often disguises true target proteins because the signals from target proteins can be overwhelmed by those from abundant sticky proteins.¹⁷

To overcome this limitation, a photoaffinity-based approach has been considered for successful target ID.⁶ This approach employs a photoactivatable moiety that can generate reactive species upon ultraviolet (UV) irradiation and create covalent bonds with adjacent molecules. Therefore, covalent linkages to target proteins can be established if the target ID probe contains a photoactive moiety, which allows for analysis by various methods because the readout is no longer affected by experimental conditions or subsequent washing steps. Owing to this covalent crosslinking of the target ID probe with proteomes, it

becomes conceptually possible to pursue target ID of bioactive small molecules with low efficacy (even micromolar activity),¹⁸ whereas the conventional affinity-based target ID approach requires that the probe have a single-digit nanomolar activity or better to ensure successful target ID via its high affinity toward target proteins. Therefore, photoaffinity-based target ID can significantly accelerate phenotype-based drug discovery through the target identification of early-stage hit compounds. Photoaffinity-based target ID can also be applied for identifying off-target proteins of FDA-approved drugs, which was previously difficult due to the low affinity between off-target proteins and the small molecules.^{19,20}

Despite these successful results, the major obstacle for photoaffinity-based target ID is the nonspecific labeling of sticky proteins.²¹ Even though photoaffinity-based target ID established a new paradigm in the target ID process, covalent labeling of non-target proteins by photoactivatable moieties often misguides researchers by generating a massive list of potential binding proteins especially when the probes' activity is relatively low. Nonspecific interactions between probes and other proteins are driven by the hydrophobic effect and electrostatic interaction, which often hides the target protein.²²

Several research groups have been studying the effect of photoaffinity linker (PL) to tackle this issue. Yao et al. rationally designed a series of PLs for the preparation of target ID probes of known bioactive small molecules to enhance the efficiency of the target ID process.^{13,14} Minimization of PLs helps to increase the number of probe binding proteins. However, the functional perturbation of identified proteins by bioactive small molecules was not

confirmed. Cravatt et al. investigated the protein interaction map of PL-embedded small molecules,^{23,24} but the scope was limited to small molecule libraries originally designed with benzophenone or diazirine. At the same time, Li et al. has been focusing on a DNA-templated photo-crosslinking approach, called DNA-programmed photoaffinity labeling (DPAL), to reduce the nonspecific binding of PLs. Although they reduced the nonspecific binding events, the experimental condition is only limited to cell lysate and requires increased salt concentrations in the buffer.²⁵ We recently developed a novel target ID technology termed fluorescence in two-dimensional gel electrophoresis (FITGE) to differentiate target proteins from nonspecific binding proteins.^{18,26,27} Even with these tremendous efforts, nonspecific labeling is still a big hurdle in photoaffinity-based target ID and needed to be investigated in detail.²⁸⁻³⁰ Only the labeling efficiency of photoactivable moieties has been conducted and compared rather than nonspecific bindings of those.³¹ To the best of our knowledge, no systematic study about the protein-labeling profiles of photoaffinity linkers has been performed. Electrophile linkers for activity-based protein profiling (ABPP) rather than PLs were only investigated by Weerapana et al. for nonspecific proteome labeling analysis of target ID probes.³² To decipher nonspecific binding events of PLs, we decided to investigate PL-dependent binding proteins and compare their protein lists.

Herein, we report the systematic study of the protein-binding profiles of photoaffinity linkers containing three different photoactivatable moieties—benzophenone (BP), diazirine (DA), and arylazide (Az). In contrast to the previous conception, PLs label different sets of proteins in a structure-

dependent manner. Once we understand the general binding patterns of each PL, we can use this finding to design target ID probes and to efficiently eliminate non-target proteins from a list of potential target proteins. We also observed the PL-dependent target protein labeling event of target ID probes in 1D gel and FITGE based target ID (Figure 2.2.1). Target ID probes with at least two kinds of PLs were found to be essential for successful deconvolution of target proteins from the entire proteome.

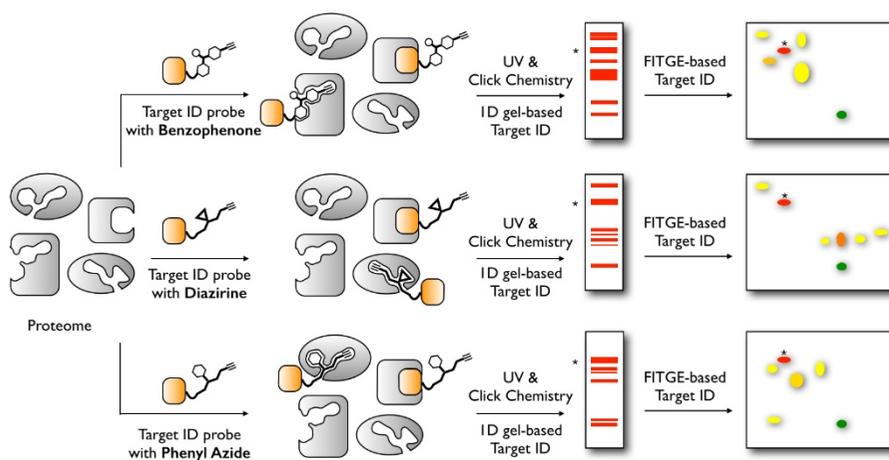


Figure 2.2.1. Schematic of photoaffinity-based target identification (ID) with different photoactivatable linkers. Each target ID probe containing a photoactivatable moiety (benzophenone, diazirine, or arylazide) can bind to a specific set of proteins in a structure-dependent manner. Additionally, non-target proteins can be specifically labeled by target ID probes. In 1D gel analysis, target proteins (*) might be masked by the large amount of nonspecific binding and specific binding of the photoaffinity linker. In FITGE-based target ID, target proteins (*) can be distinguished from nonspecifically bound proteins and those specifically bound to photoaffinity linkers. Systematic study of the protein-labeling profiles of photoaffinity linkers allows efficient identification of the target proteins (*) by eliminating non-target proteins (orange spots in the 2D gel).

2.2.3. Results and discussion

Proteome-labeling patterns are PL structure-dependent.

To investigate the protein-binding profiles of photoaffinity linkers, we designed and synthesized photoaffinity linkers **1**, **2**, and **3**, each with a frequently used photoactivatable moiety (BP, DA, and Az, respectively) and an acetylene moiety for fluorescent labeling via click chemistry (Figure 2.2.2A and Supporting Information). Even though with easy synthetic scheme and low reactivity to water, BP requires long irradiation time to generate a triplet diradical, which might increase the nonspecific protein binding. DA generates the carbene intermediate upon UV irradiation. Due to high reactivity and short lifetime, DA can rapidly form a covalent bond with target proteins. But, this high reactivity of carbene often leads to the easy quenching in water, which might decrease the photoaffinity labeling efficiency. Upon UV irradiation, Az generates nitrene species, which have low labeling ability not only due to low reactivity of nitrene itself, but also the subsequent rearrangement of nitrene to benazirines and dehydroazepines/ketenimines. Thiol-mediated azide reduction to amine is another reason for low labeling efficiency. However, commercial availability and easy synthetic procedure is a key advantage of Az.²²

We treated live cells (Figure 2.2.2B) from various origins, including HeLa cells (human cervical cancer cell line), BV2 cells (mouse microglial cell line), and 293T cells (human epithelial kidney cell line) with each photoaffinity linker. The resulting cells were subjected to UV irradiation to generate reactive species (carbon radical, carbene, and nitrene, respectively) capable of forming covalent

bonds with nearby proteins. Proteins crosslinked to each PL were then labeled with Cy5-azide via copper-mediated bioorthogonal click chemistry and visualized by fluorescence gel scanner after 1D gel electrophoresis.

Each PL showed different labeling patterns in all cell lines (Figure 2.2.2B). Interestingly, the reported labeling efficiency of the photoactivatable groups (labeling efficiency: Az > DA > BP)³¹ does not match with the fluorescent intensity of protein bands labeled by PLs. In contrast, the proteome-labeling patterns were dependent on the identity of the photoaffinity linker. The BP-based PL (**1**) labels a number of proteins, especially those with having 40–70-kDa molecular weight; the DA-based PL (**2**) specifically labels a ~30-kDa protein in all cell lines; the labeling pattern of the Az-based PL (**3**) appears to be the nonspecific but is very different from those of **1** and **2**. These findings indicate that PLs specifically bind to certain proteins in living cells and form covalent crosslinks in a structure-dependent manner.

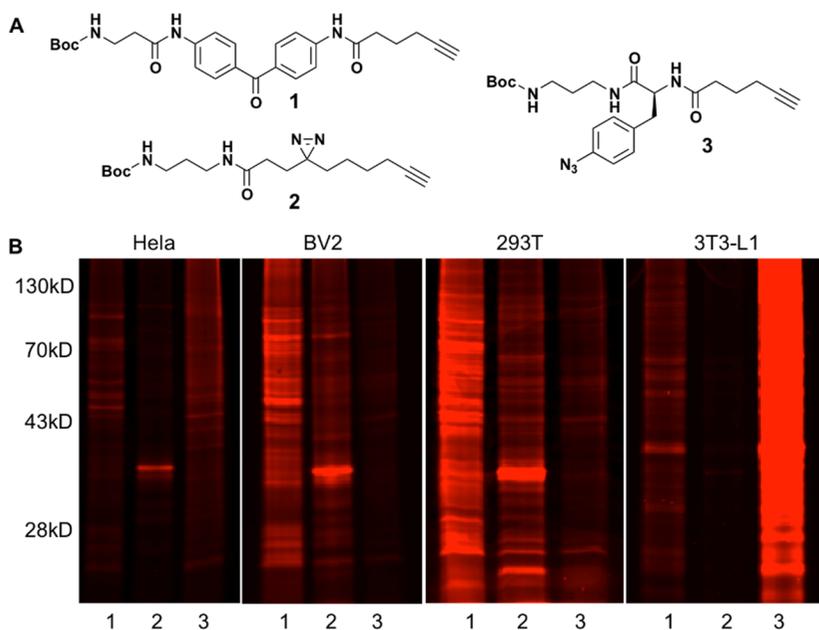


Figure 2.2.2. A) Structure of three photoactivatable moieties with acetylene moiety for click chemistry. B) Photoaffinity linkers **1**, **2**, and **3** (10 μ M each) were used to treat live HeLa cells (human cervical cancer cell line), BV2 cells (mouse microglial cell line), and 293T cells (human epithelial kidney cell line). After UV irradiation, the cells were lysed, and the cell lysates were labeled with Cy5-azide.

However, it was difficult to visualize the labeling pattern of the entire proteome using 1D gel electrophoresis due to its limited resolution. Thus, we applied 2D gel electrophoresis to identify the proteome-labeling profiles of each PL. Briefly, the proteomes of living HeLa, BV2, and 293T cells were UV-crosslinked with **1**, **2**, and **3**, and then linked to Cy5-azide by bioorthogonal click chemistry for visualization. The resulting proteomes were analyzed with 2D gel electrophoresis. The proteome-labeling pattern in the 2D gels showed distinct differences for each PL (Figure 2.2.3). Although the proteome-labeling patterns of **1** and **3** seemed nonspecific and random in the 1D gel, the higher resolution of 2D gel electrophoresis revealed that the proteome-labeling pattern

of each PL was not random but rather structure-dependent throughout the three cell lines. We analyzed the proteome-labeling pattern of each PL in the 2D gels; **1** specifically labeled proteins with isoelectric point (pI) 3–4 and 50–70-kDa in all tested cell lines (Figure 2.2.3, left column). Compared to the proteomes labeled by **2** and **3**, the upper left region of the 2D gel was intensively labeled by **1**. In contrast, **2** specifically labeled proteins with pI 7–10 and 35-kDa in all tested cell lines (Figure 2.2.3, center column), which as a result allows these proteins to be used as a signature for **2**. Lastly, **3** labeled proteins with pI 3–4 and 50–70-kDa in all tested cell lines (Figure 2.2.3, right column).

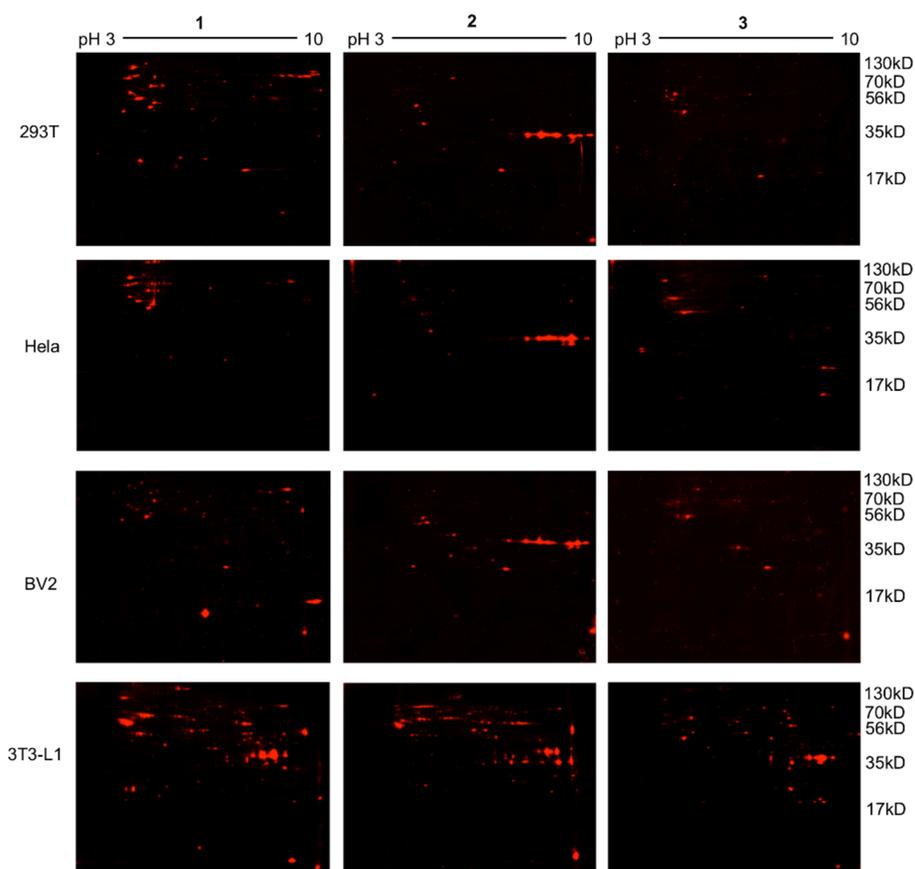


Figure 2.2.3. Proteome-labeling pattern of each photoaffinity linker analyzed by 2D gel

electrophoresis. 10 μ M Photoaffinity linkers **1** (left), **2** (center), and **3** (right) were used to treat live 293T (top), HeLa (center), and BV2 (bottom) cells. After UV irradiation, the cells were lysed, and the lysates were labeled with Cy5-azide via bioorthogonal click chemistry. The proteome-labeling pattern of each photoaffinity linker was analyzed by 2D gel electrophoresis.

MS analysis shows PL-specific protein labeling.

Although the proteins labeled by **1** and **3** were both located at the upper left region of the 2D gel, the individual proteins labeled by each PL were strikingly different. The fluorescence intensity and number of proteins labeled by **1** were higher than those for **3**. Based on the general proteome-labeling patterns we observed, we pursued the identification of individual proteins specifically labeled by each PL. For this profiling, we excised all the proteins specifically labeled by **1**, **2**, and **3** and performed mass spectrometry (MS) analysis of more than 300 protein spots via in-gel digestion. Of the hundreds of proteins labeled by each PL, those identified in at least two different cell lines as well as being confirmed by MS/MS analysis are listed in Table 2.2.1. The full list of identified proteins is given in the Supplementary Information.

Vimentin, β -globin, tubulin, DEAD box polypeptide, heterogeneous nuclear ribonucleoprotein M, polyadenylate-binding protein, mitochondrial ATP synthase, heat shock protein 90 (HSP90), and heat shock protein 60 (HSP60) were frequently labeled by **1**, among which vimentin, β -globin, tubulin, DEAD box polypeptide, polyadenylate-binding protein, and heterogeneous nuclear ribonucleoprotein M were identified in all tested cell lines. **2** showed specific binding to glyceraldehyde-3-phosphate dehydrogenase (GAPDH), lactate dehydrogenase (LDH), mitochondrial malate dehydrogenase

2 (MDH2), RPS2 protein, macropain, and PKM2 protein. The ~30-kDa LDH is a signature protein specifically labeled by **2**, and LDH, GAPDH, and mitochondrial MDH2 were identified in all tested cell lines. Overall, **2** showed much weaker protein labeling than the others, especially at the upper left region of the 2D gel. **3** specifically labeled HSP60, tubulin, dermcidin, FLJ00410, perilipin, as well as actin which was identified in all three tested cell lines. Among those proteins specifically or commonly labeled by PLs, tubulin and actin were all found to be labeled by three PLs (Table S2.2.1). β -globin is labeled by **1** and **2**, HSP60 is labeled by **1** and **3**, and dermcidin protein is labeled by **2** and **3**. 23 out of 27, 15 out of 19, and 4 out of 8 proteins are specifically labeled by **1**, **2**, and **3**, respectively. The location of protein labeled by each PL is also structure dependent. The proteins labeled by **1** and **2** are located in cytoplasm, nucleus, and mitochondria. About 40 percentage of proteins in the list are located in cytoplasm, 20~30 percentage of proteins in nucleus, and 10 percentage of proteins in mitochondria. The proteins labeled by **3** are dominantly located in cytoplasm in contrast to those labeled by **1** and **2**. As can be seen, the majority of the identified proteins are specifically labeled by PLs in a structure-dependent manner. Based on these results, we revealed the different protein-crosslinking profiles of PLs containing three different photoactivatable moieties.

Photoaffinity linker	Protein name	BV2	HeLa	293T	Protein Location
1	vimentin	3	2	3	Cytoplasm
	beta-globin	1	2	4	Intracellular
	tubulin	3	1	1	Cytoplasm
	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	1	2	2	Nucleus
	polyadenylate-binding protein	1	2	1	Cytoplasm
	heterogeneous nuclear ribonucleoprotein M	1	1	2	Nucleus
	HSP90	3		4	Cytoplasm
	HSP60	1		4	Cytoplasm
	YTH domain family protein	2		2	Nucleus, Cytoplasm
	RCC2 protein		1	3	Nucleus
	actin	2	1		Cytoplasm
	alpha-hemoglobin	2		1	Intracellular
	HSP70		1	2	Cytoplasm
	mitochondrial ATP synthase	2		1	Mitochondria
	Pspc1NONO		1	2	Nucleus
	titin		1	2	Membrane
	poly(A) binding protein		1	2	Cytoplasm
	CAP, adenylate cyclase-associated protein 1	1		1	Cytoplasm
	eukaryotic translation initiation factor 4A		1	1	Nucleus, Cytoplasm
	enoyl-Coenzyme A hydratase	1		1	Mitochondria
	lamin A/C transcript variant 1		1	1	Cytoplasm
	NADH dehydrogenase (ubiquinone) flavoprotein 3		1	1	Mitochondria
	KIAA1150 protein	1		1	Intracellular
	splicing factor 3a	1		1	Nucleus
	TNF receptor-associated protein 1		1	1	Mitochondria
elongation factor 1 alpha		1	1	Intracellular	
Chain C, Structure Of The Ran-Gppnhp-Ranbd1 Complex		1	1	Intracellular	
2	L-lactate dehydrogenase	4	2	1	Cytoplasm
	Glyceraldehyde-3-phosphate dehydrogenase	2	3	2	Cytoplasm
	mitochondrial malate dehydrogenase 2	1	1	1	Mitochondria
	beta-globin		3	3	Intracellular
	heterogeneous nuclear ribonucleoprotein A2/B1		3	2	Nucleus
	Tubulin		2	2	Cytoplasm
	RPS2 protein		2	1	Cytoplasm
	HNRPA1 protein		2	1	Cytoplasm
	Porin 31HM		2	1	Cytoplasm
	ribosomal protein S2		1	2	Cytoplasm
	dermcidin preproprotein		1	2	Intracellular
	macropain		1	2	Nucleus, Cytoplasm
	actin	1	1		Cytoplasm
	triosephosphate isomerase	1		1	Nucleus
	Voltage-dependent anion-selective channel protein 3		1	1	Mitochondria
	nitric oxide synthase interacting protein	1	1		Nucleus
	PKM2 protein		1	1	Cytoplasm
	mitochondrial ribosomal protein L4	1		1	Nucleoli
ankyrin repeat domain 23		1	1	Nucleus	
3	Actin	2	1	1	Cytoplasm
	HSP60		2	1	Cytoplasm
	tubulin		1	2	Cytoplasm
	dermcidin		1	2	Intracellular
	FLJ00410		1	1	Intracellular
	Perilipin-3		1	1	Cytoplasm
	PAWR		1	1	Membrane, Cytoplasm
RAN binding protein 1		1	1	Intracellular	

^aIdentified proteins were colored in red: BV2, blue: HeLa, and green: 293T. Numbers indicate the identification frequency of each protein in the cell lines. The cellular location of PL binding proteins is colored. Full list of linker-binding proteins are available in Supporting Information

Table 2.2.1. The signature proteins labeled by photoaffinity linkers 1, 2, and 3 in three different live cell lines identified by MS analysis of protein spots.^a

Proteins labeled by target ID probes are PL-structure dependent.

We then applied the protein-labeling pattern of each PL to improve the efficiency in the target identification of bioactive small molecules. From the PL-dependent protein labeling data, we postulated that the proteins labeled by target ID probes would differ on the basis of PLs used. To observe the PL-dependent protein labeling, we designed a set of target ID probes containing three different PLs (1–3) for three bioactive molecules; a neuroinflammation inhibitor **ICM**,²⁷ a glucose uptake enhancer **PPAR-e**,¹⁸ and an anti-proliferative agent **Tubulin-i**²⁶ discovered from our in-house 4000-member compound library constructed by a privileged substructure-based diversity-oriented synthesis (pDOS) strategy.^{33,34} Along with our reported two target ID probes, Tubulin-BP²⁶ and ICM-BP²⁷, we newly synthesized seven target ID probes [Tubulin-DA, -Az; PPAR-BP, -DA, -Az; and ICM-DA, -Az] to complete the set of target ID probes with three different PLs (Figure 2.2.4A and Supporting Information). Especially, PPAR-BP, -DA, and -Az were designed using new PPAR γ agonist **PPAR-e** with improved transcriptional activity (EC₅₀: 3 nM; Figure S2.2.3), compared to our originally reported PPAR γ agonist P29C06 (EC₅₀: 1500 nM).¹⁸ Prior to the target ID, we confirmed the comparable cellular activities of these nine target ID probes to their original molecules, **ICM**, **PPAR-e**, and **Tubulin-i**. (Figure S2.2.1–S2.2.4). We then pursued photoaffinity-based target ID with these target ID probes in the absence or presence of soluble competitors (**ICM**, **PPAR-e**, and **Tubulin-i**). From the resulting proteome in 1D gel (Figure 2.2.4B), we clearly observed the meaningful PL-dependent protein-labeling pattern shift of target ID probes in

all three cases. DA-based target ID probes (ICM-DA, PPAR-DA, and Tubulin-DA) labeled the ~30-kDa proteins, the signature proteins of **2**. The protein-labeling pattern of Az-based target ID probes (ICM-Az, PPAR-Az, and Tubulin-Az) was slightly different from that of the BP-based target ID probes (ICM-BP, PPAR-BP, and Tubulin-BP).

Target protein-labeling of target ID probes is independent to the type of PLs in 1D gel electrophoresis. ICM-derived target ID probes (ICM-BP, ICM-DA and ICM-Az) can label HMGBs (Figure 2.2.4B left panel, asterisks in lanes 1, 3, and 5, respectively). However, even with improved cellular activities of **PPAR-e**-derived target ID probes (PPAR-BP, -DA, and -Az), we failed to identify the target protein, PPAR γ , on the 1D gel, which is consistent with our previous study.¹⁸ In addition, **Tubulin-i**-derived target ID probes (Tubulin-BP, -DA, and -Az) failed to label the target protein, Tubulin.²⁶ For an in depth analysis of protein-labeling patterns of nine target ID probes, we performed FITGE-based target ID.²⁶ Briefly, cellular proteins that interact with target ID probes were captured via covalent bonds upon UV irradiation. The labeled proteomes in the absence or presence of **ICM**, **PPAR-e**, and **Tubulin-i** were visualized with Cy5-azide and Cy3-azide, respectively. Both samples were then pooled and analyzed together using 2D gel electrophoresis. After scanning for both Cy5 and Cy3 channels, the gel images were merged to show distinct and overlapped protein spots (Figure 2.2.5; red: Cy5, green: Cy3, and yellow: Cy5 and Cy3) for visualization of specific (red) and nonspecific (green and yellow) labeling.

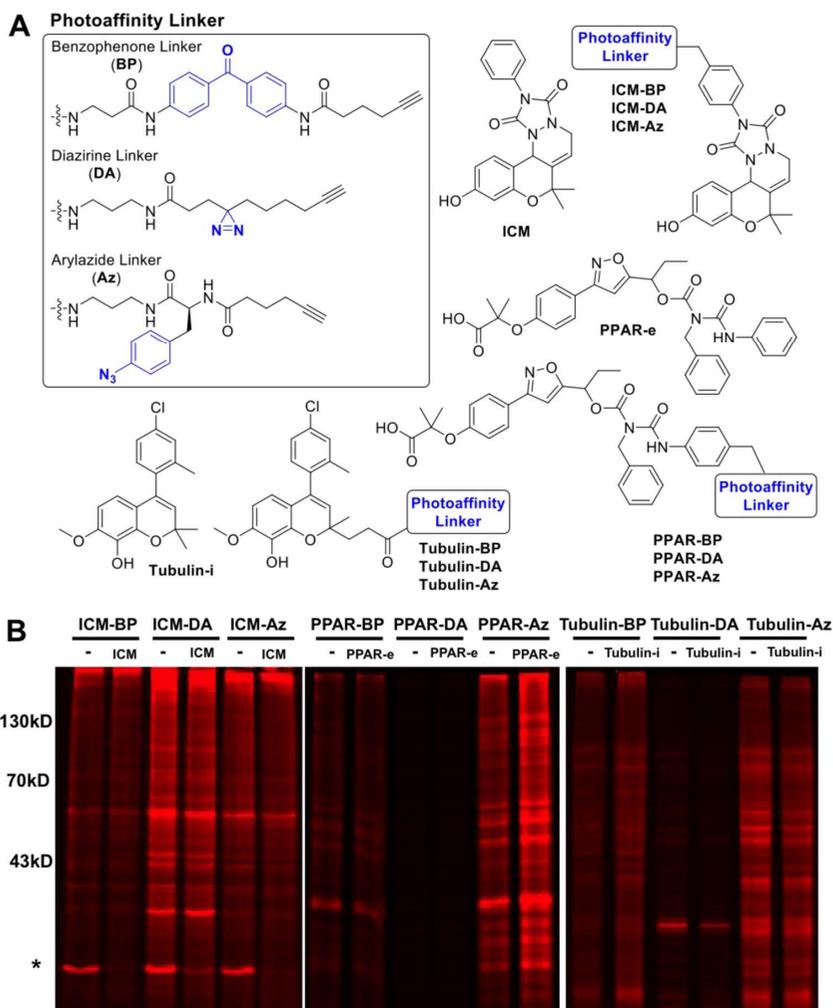


Figure 2.2.4. A) Structure of nine target ID probes of **ICM**, **PPAR-e**, and **Tubulin-i** with three different photoaffinity linkers. B) Protein labeled by the target ID probes of **ICM** (10 μ M, left panel), the target ID probes of **PPAR-e** (10 μ M, center panel), the target ID probes of **Tubulin-i** (10 μ M, right panel) in the presence of **ICM**, **PPAR-e**, and **Tubulin-i** (40 μ M) was displayed on 1D gel. Photoaffinity linker-dependent protein labeling was observed in all three target ID cases. (asterisk on left panel indicates target protein of **ICM**, HMGBs)

The protein-labeling patterns for the nine target ID probes differed partially

on FITGE-based 2D gels in a PL-dependent manner. Proteins in the upper left region of the 2D gel were extensively labeled by BP-based target ID probes (Figure 2.2.5, left column). DA-based target ID probes showed significant labeling of ~30-kDa signature proteins in the central region of the 2D gel (Figure 2.2.5, center column). Compared to BP-based probes, while the overall labeling intensity by DA-based probes was less intense in the 1D gel, we did observe clear differences in PL-dependent labeling patterns of target ID probes in the FITGE-based target ID (Figure 2.2.5). The labeling efficiency of Az-based probes was less intense in the upper left region of the 2D gel than that of BP-based probes (Figure 2.2.5, right column), which is consistent with that of PLs **1** and **3** (Figure 2.2.3). From these results, we could observe that the protein-labeling pattern shift by target ID probe is PL-structure dependent. MS analysis of yellow or orange protein spots showed that most of the signature proteins labeled by **1**, **2**, and **3** were labeled by target ID probes with each PL, respectively (Fig. S2.2.12–S2.2.20 and full MS data in Supporting Information).

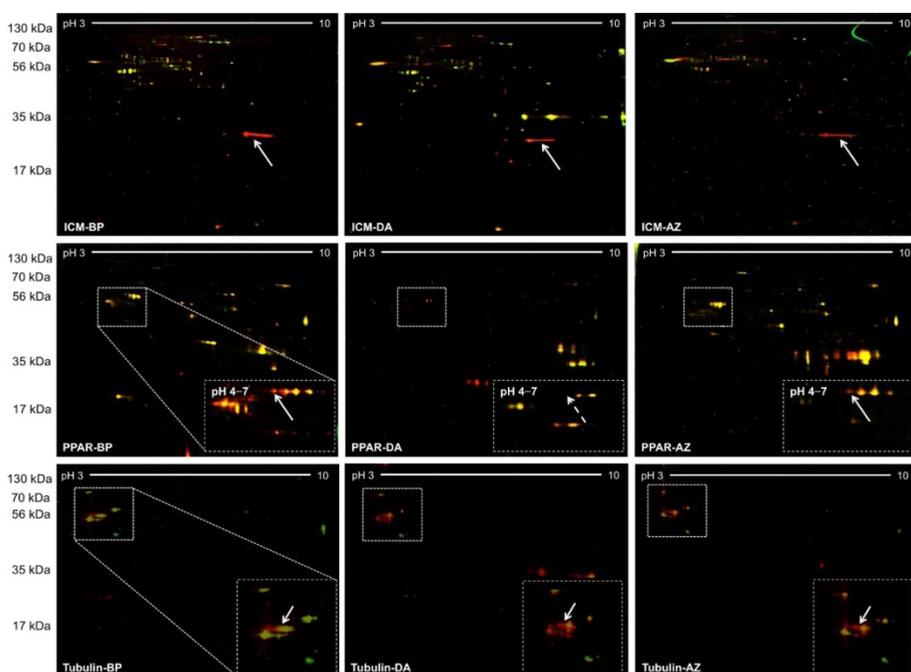


Figure 2.2.5. FITGE-based target ID using target ID probes embedded with three photoaffinity linkers. Proteomes were labeled with target ID probes (10 μ M) in the absence or presence of initial hit compounds, **ICM**, **PPAR-a**, and **Tubulin-i** (40 μ M). The labeled proteomes in the absence and presence of **ICM**, **PPAR-a**, and **Tubulin-i** were visualized with Cy5-azide and Cy3-azide, respectively. The lysates were pooled and analyzed by 2D gel electrophoresis (pH 3–10 strip or pH 4–7 strip for isoelectric focusing). The gel was scanned in the Cy5 and Cy3 channels using a fluorescence scanner, and the image were merged for analysis. Red: Cy5, green: Cy3, yellow: Cy5 and Cy3. White arrows indicate the target proteins of **ICM**, **PPAR-e**, and **Tubulin-i**.

Efficient deconvolution of target protein using PL-specific binding protein list

ICM-BP, ICM-DA, and ICM-Az exhibited specific binding to HMGBs, the confirmed target protein of **ICM** (Figure 2.2.5, first row, white arrow), as in 1D

gel.²⁷ Notably, the labeling pattern of HMGBs was not affected by the identity of PL in the target ID probes. The further MS analysis of additional red spots revealed that HSP90 was specifically labeled by ICM-BP, and HSP60 and actin were labeled by ICM-Az (Figure S2.2.12–S2.2.14 and Table S2.2.2). Based on the protein-labeling patterns of **1**, **2**, and **3** (Table 2.2.1), these binding proteins to ICM-derived target ID probes can be efficiently ruled out from the candidate list of target proteins. In the case of **PPAR-e**-derived target ID probes, we observed other interesting protein spots in red at the center of the 2D gel along with PPAR γ (Figure 2.2.5, second row). MS analysis showed that eukaryotic translation elongation factor 1 alpha 1, β -globin, and enoyl-Coenzyme A hydratase were specifically labeled by PPAR-BP. In contrast, macropain, mitochondrial MDH2, triosephosphate isomerase, and GAPDH were labeled in red by PPAR-DA. However, those proteins are the signature proteins labeled by **1**, **2**, and **3** (Table S2). Therefore, those proteins could be eliminated from the candidate list of target proteins. Along with target protein of **Tubulin-i** (tubulin), several red spots were observed in FITGE-based target ID using Tubulin-BP and Tubulin-DA (Figure 2.2.5, third row). MS analysis gave that HSP90 and actin were also labeled in red at the lower left region of the 2D gel by Tubulin-BP. Protein spots at the ~30-kDa region were specifically labeled by Tubulin-DA and identified as LDH, GAPDH, mitochondrial MDH2, Porin 31HM, heterogeneous nuclear ribonucleoprotein A2/B1, ribosomal protein S2, and voltage-dependent anion-selective channel protein, that are the signature proteins of **2** (Table S2.2.2).

If not for our protein-binding analysis of PLs, we might have wrongly

selected these proteins as potential target proteins of **Tubulin-i**. Hence, our systemic protein-labeling analysis of PLs saved us a lot of time and efforts in the subsequent target confirmation. In an ideal situation for the competition experiment of target ID probes with active ligands, only target proteins should be labeled in red. However, soluble competitors having low activity (micromolar activity) often fail to compete out target ID probes only from target protein as we could observe in the all case of ICM, PPAR, and Tubulin target ID. We believe that when the concentration of soluble competitors is increased high enough (40 μ M in this work) to compete out target ID probes from target proteins, soluble competitor also starts to bind non-target proteins nonspecifically by hydrophobic interactions.²² Therefore, the nonspecific interactions between target ID probes and non-target proteins, driven by hydrophobic interactions, would be interfered by high concentration of soluble competitors. In this context, our investigation of PL-specific binding proteins could help to deconvolute target protein efficiently. Although we could not eliminate all non-target proteins shown in red, the protein-labeling patterns of PLs helped us to narrow down the target protein candidates in the identified protein list of nine target ID cases described above.

Target protein-labeling efficiency by target ID probes is PL-dependent.

During this study, we observed another interesting feature in the case of **PPAR-e**- and **Tubulin-i**-derived target ID probes. Even though the cellular activity of the target ID probes was conserved, the proteinlabeling efficiency was affected by PLs embedded in the target ID probes. To ensure the resolution

of the targets of **PPAR-e**,¹⁸ we narrowed the pH range from 3–10 to 4–7 for isoelectric focusing. We were thus able to identify the target protein of **PPAR-e**, PPAR γ , for both PPAR-BP and PPAR-Az (Figure 2.2.5, second row left and right panel, white arrows). However, we failed to identify PPAR γ using PPAR-DA (Figure 2.2.5, second row center panel, white dashed arrow), which indicates that the structure of the target ID probe can negatively affect the outcome of target ID process despite the conserved cellular activity of the probe. Interestingly, the target protein-labeling pattern is not associated with the labeling efficiency of PLs. Considering the PL labeling efficiency (Az > DA > BP), Az and DA might be the good choice for target protein labeling, but the target ID results showed DA was not a good choice as a PL of target ID probe.

In the case of the target protein of **Tubulin-i**, tubulin is located inside the dashed rectangle in the upper left region of the 2D gel (Figure 2.2.5, third row). Enhancement of this region also revealed distinct target protein-labeling patterns of each target ID probe. Notably, tubulin was labeled in yellow as a nonspecific binding protein for Tubulin-BP, which is consistent with our previous report²⁶ (Figure 2.2.5, third row, left inset, white arrow). However, tubulin was successfully labeled as a specific target in the case of both Tubulin-DA and Tubulin-Az (Figure 2.2.5, third row, middle and right inset, white arrow). Simple changes of photoactivatable moieties in the target ID probe drastically affected the outcome of the photoaffinity-based target ID approach. This result explains our initial failure in target ID with Tubulin-BP via competition with soluble ligand **Tubulin-i**. Actually, the BP moiety itself was previously reported to bind tubulin,³⁵ which we also confirmed with our

protein-labeling profile study of three different PLs (Table 2.2.1). Due to the fact that both **Tubulin-i** and BP can bind tubulin, Tubulin-BP has two tubulin-binding elements and consequently enhanced binding affinity for tubulin. Therefore, the 4-fold excess of soluble ligand **Tubulin-i** was not able to compete out the binding of Tubulin-BP toward tubulin in photoaffinity-based target ID, which disguised tubulin in yellow as a nonspecific binding protein. When we changed the photoactivatable moiety from BP to DA or Az, the synergistic binding affinity of two tubulin-binding elements of Tubulin-BP was vanished. The resulting target ID probes, Tubulin-DA and Tubulin-Az, could be competed out by soluble ligand **Tubulin-i**, which enabled the identification of tubulin in red as a specific binding protein.

This systematic PL-dependent protein-labeling study with **PPAR-e** and **Tubulin-i** emphasizes the importance of the choice of photoaffinity linker in the design of target ID probes. Even when the activities of the target ID probes are conserved, the labeling efficiency for target proteins can be affected by the PL structures in target ID probes. Therefore, synthesis of target ID probes with various PLs is crucial for successful target identification. By using target ID probes with three PLs, we could identify the target protein of **PPAR-e** and **Tubulin-i**. Had we pursued the target ID using a single type of target ID probe, we might have missed the real target protein from the candidate proteins. Therefore, the preparation of target ID probes with at least two kinds of PLs is essential for the successful deconvolution of target proteins from the entire proteome.

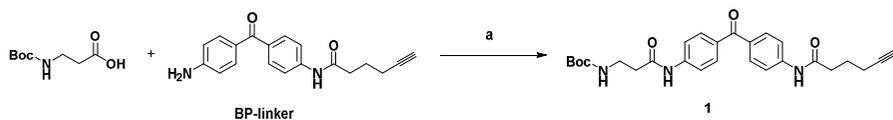
2.2.4. Conclusion

In this study, we demonstrated the structure-dependent protein labeling of photoaffinity linkers for the first time. Our systematic protein-labeling pattern analysis of PLs can facilitate the discrimination of a target protein from specific non-target proteins that we encounter in the actual target ID process. We also demonstrated the importance of PL in the design of target ID probes using three case studies. These studies showed that target proteins might be missed if the proper PL were not chosen for the target ID probe. Although we could not suggest the best PL, our data shows that at least two types of PLs are essential for successful target ID process. Furthermore, the mechanistic understanding of PL effects in photoaffinity-based target ID will be investigated in the upcoming research. Therefore, our systematic understanding of PL-dependent protein labeling is critical to eliminate non-specific binding proteins from target candidates, thereby increasing the success rate of the photoaffinity-based target ID process.

2.2.5. Experimental Section

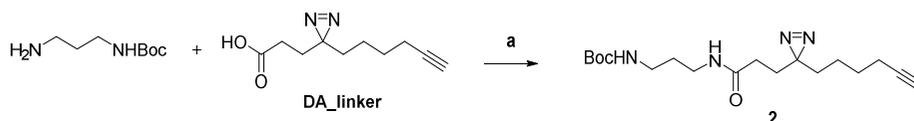
Detailed synthetic procedures of all new compounds (**1**, **2**, **3**, **3'**, BP_linker, DA_linker, Az_linker, ICM-BP, ICM-DA, ICM-Az, PPAR-BP, PPAR-DA, PPAR-Az, Tubulin-BP, Tubulin-DA, and Tubulin-Az) were described in the supporting informations of *ACS Chem. Biol.* **2016**, *11*, 44–52.

Supplementary Scheme S2.2.1. Synthesis of photoaffinity linker 1.



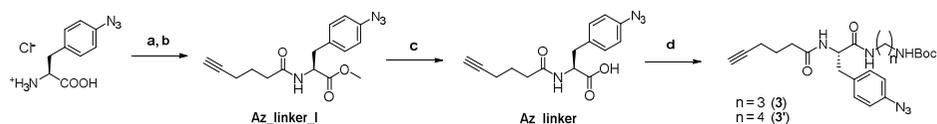
Reagents and conditions: (a) EDC, HOBT, TEA, r.t., 12 h.

Supplementary Scheme S2.2.2. Synthesis of photoaffinity linker 2.



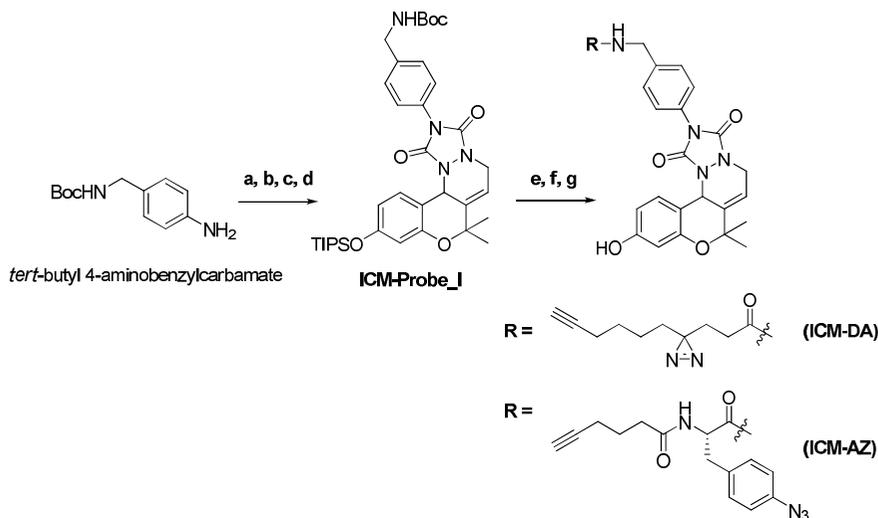
Reagents and conditions: (a) EDC, HOBT, TEA, r.t., 12 h.

Supplementary Scheme S2.2.3. Synthesis of photoaffinity linker 3 and 3'.



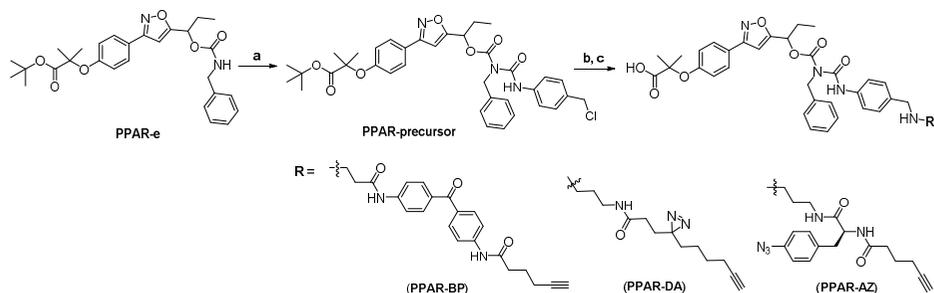
Reagents and conditions: (a) thionyl chloride, methanol, r.t., 12 h; (b) 5-hexynoic acid, oxalyl chloride, triethylamine, THF, 0 °C → r.t., 4 h; (c) lithium hydroxide, water/methanol/THF(1:1:3), r.t., 12 h; (d) EDC, HOBT, TEA, r.t., 12 h.

Supplementary Scheme S2.2.4. Synthesis of ICM target probes



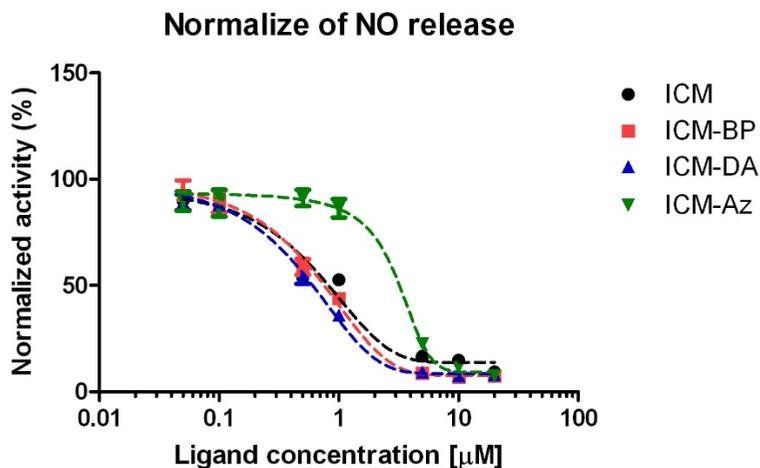
Reagents and conditions: **(a)** methyl carbazate, carbonyl diimidazole, THF, r.t. \rightarrow 70 °C, 5 h; **(b)** KOH, water/THF, 70 °C, 12 h; **(c)** iodobenzene diacetate, THF, r.t., 1 h; **(d)** **1**, THF/toluene (1:1), r.t., 12 h; **(e)** trifluoroacetic acid, DCM, r.t., 12 h; **(f)** **DA_linker** or **Az_linker**, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, hydroxybenzotriazole, triethylamine, 1,4-dioxane, r.t., 12 h; **(g)** HF/pyridine/THF(1:1:18), r.t., 6 h.

Supplementary Scheme S2.2.5. Synthesis of target ID probes of anti-diabetic agent PPAR-e

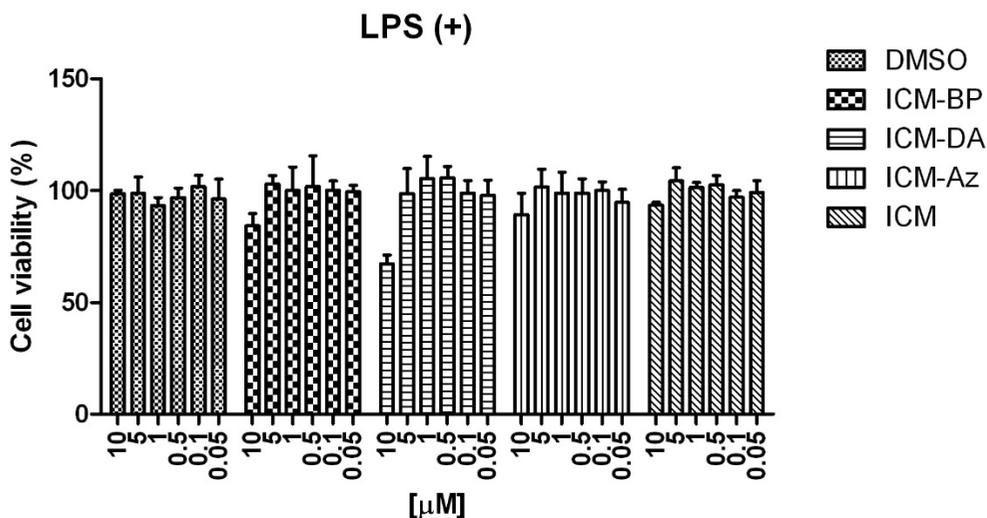


Reagents and conditions: **(a)** CuCl, 4-(Chloromethyl)phenyl isocyanate, DIPEA, DCM/THF/DMF (4:4:1), r.t., 24 h; **(b)** appropriate amine-TFA salt, DIPEA, ACN, r.t., 12 h; **(c)** trifluoroacetic acid, DCM, r.t., 6 h.

2.2.6. Supporting Information

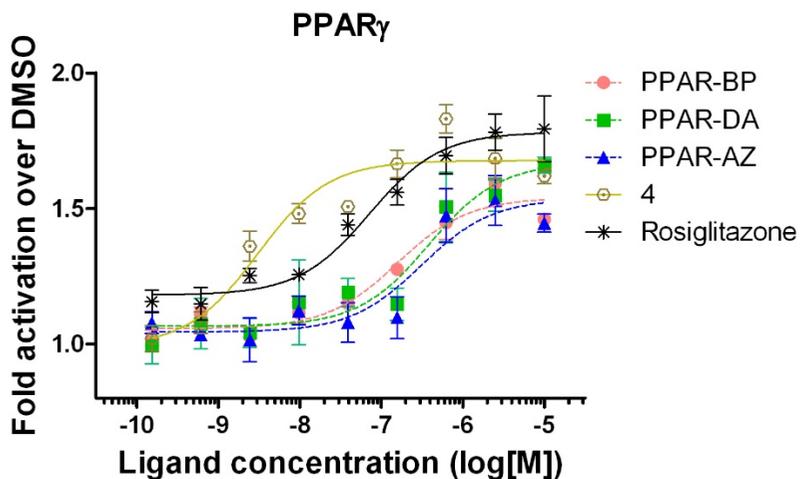


Supplementary Figure S2.2.1. Griess Assay (Nitrite release inhibition measurement) of ICM, and three ICM target ID probes upon the LPS stimulation on BV2 microglial cells.

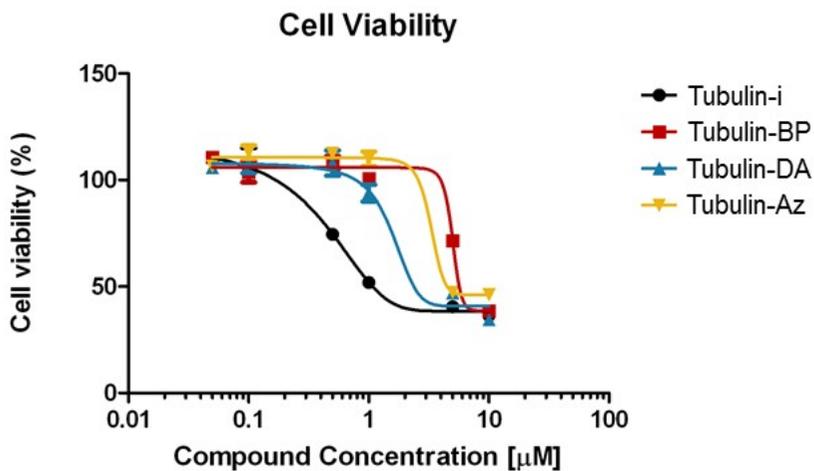


Supplementary Figure S2.2.2. Non-cytotoxic effect of ICM and ICM target ID probes on BV2 microglial cells. Cell viability was measured by WST assay after 24 h incubation in the same

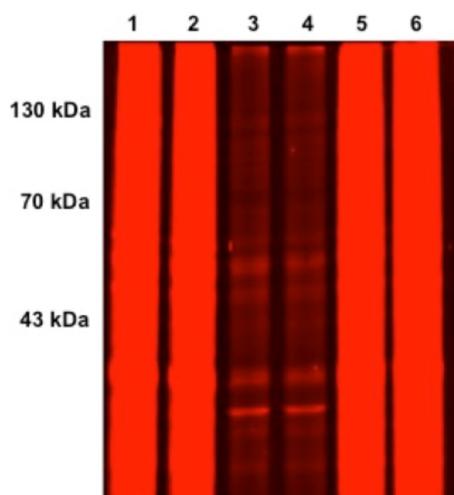
manner as Figure S2.2.2.



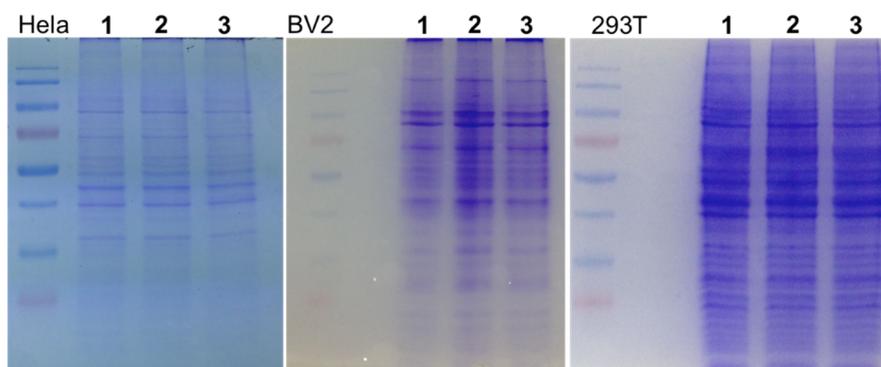
Supplementary Figure S2.2.3. Transactivation profiles of rosiglitazone, anti-diabetic agent **PPAR-e**, and three target ID probe of **PPAR-e** using a PPAR-derived reporter gene in 293T cells at 24 h after treatment.



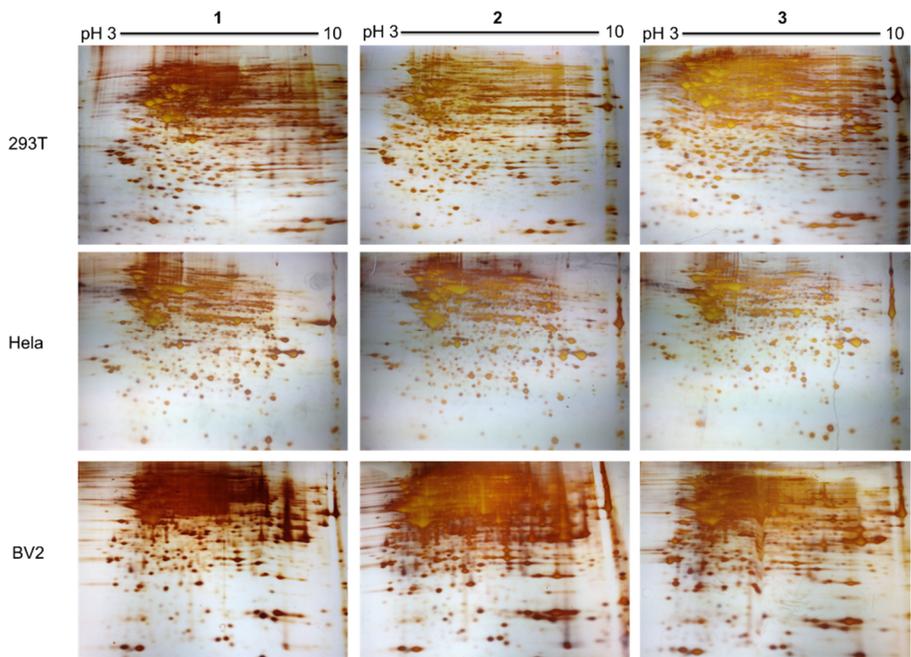
Supplementary Figure S2.2.4. Cytotoxic effect of anti-cancer agent **Tubulin-i** and three target ID probes of **Tubulin-i** on HeLa cervical carcinoma cell line.



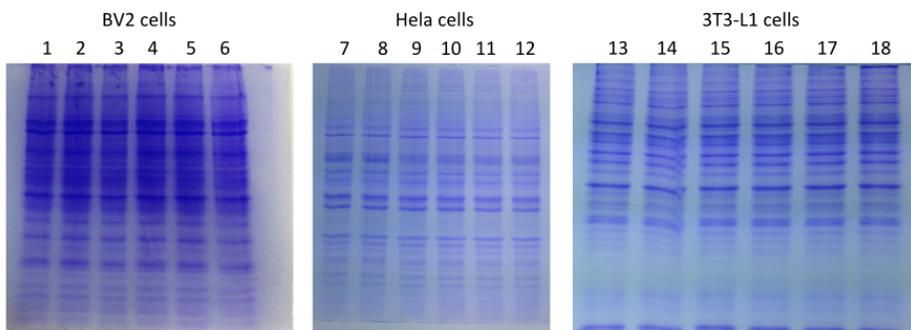
Supplementary Figure S2.2.5. Fluorescence scan imaging of Figure 2.2.4B center panel with high PMT.



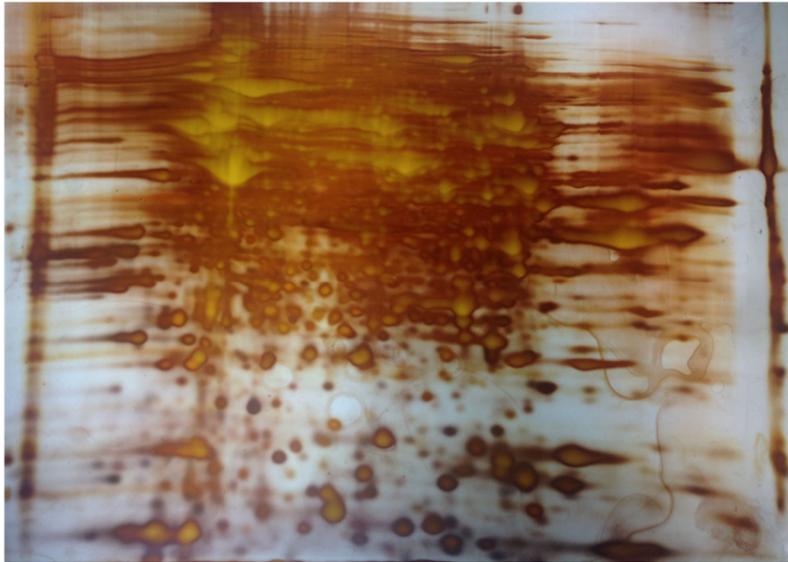
Supplementary Figure S2.2.6. Coomassie staining of 1D gel image in Figure 2.2.2B. Expression pattern of proteomes was not changed by proteome labeling by photoaffinity linker, but the proteome-labeling pattern was affected by the identity of photoaffinity linker in a structure-dependent manner.



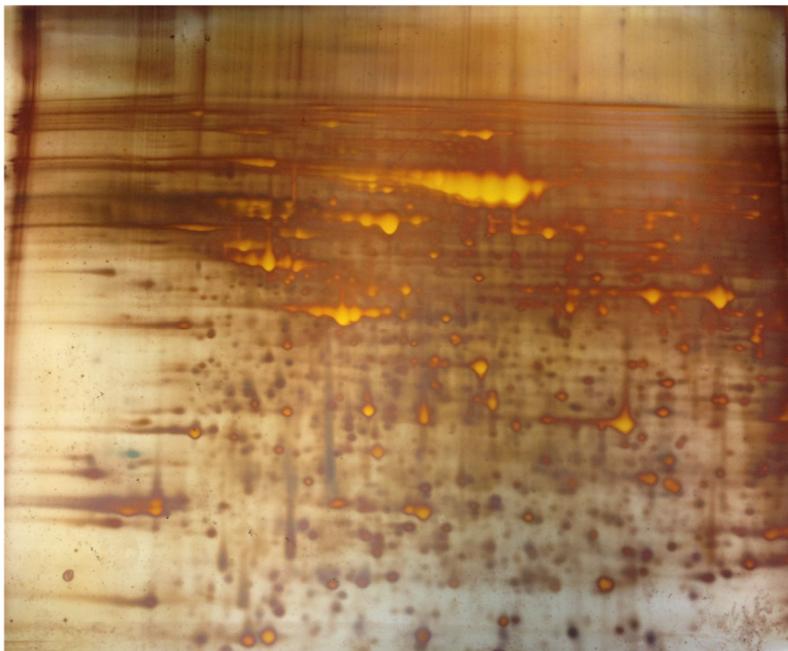
Supplementary Figure S2.2.7. Silver staining of 2D gel image in Figure 2.2.3. Whole proteome expression pattern was not affected by proteome labeling by photoaffinity linkers. Only few proteins out of whole proteome is labeled by each photoaffinity linkers.



Supplementary Figure S2.2.8. Coomassie staining of Figure 2.2.4B, 2.2.5B, and 2.2.6B. Cellular protein expression patterns were not affected by photoaffinity protein labeling of target ID probes.



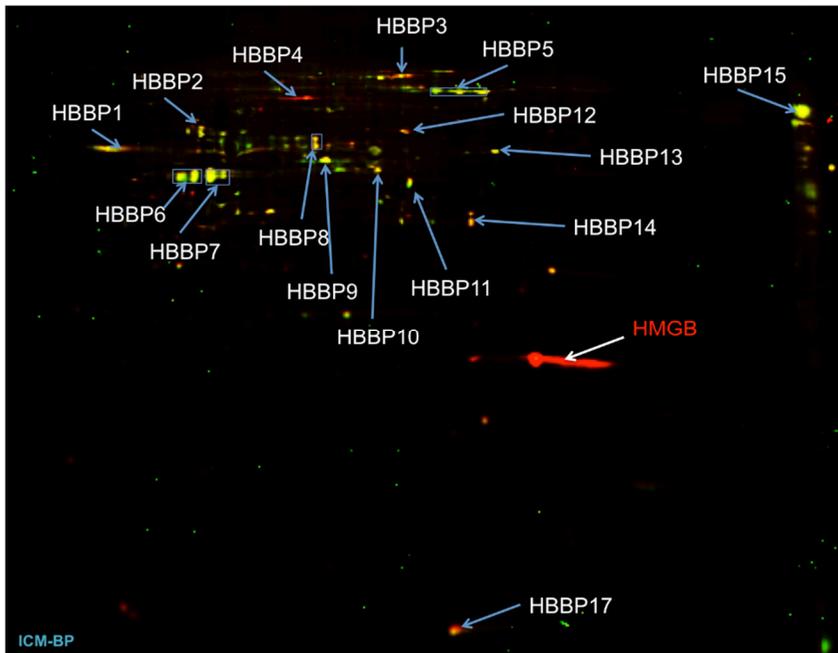
Supplementary Figure S2.2.9. Silver staining of 2D gel (pH 3–10 range) labeled by ICM-BP. Among thousand proteins, only couple of dozen proteins were labeled by the target ID probe.



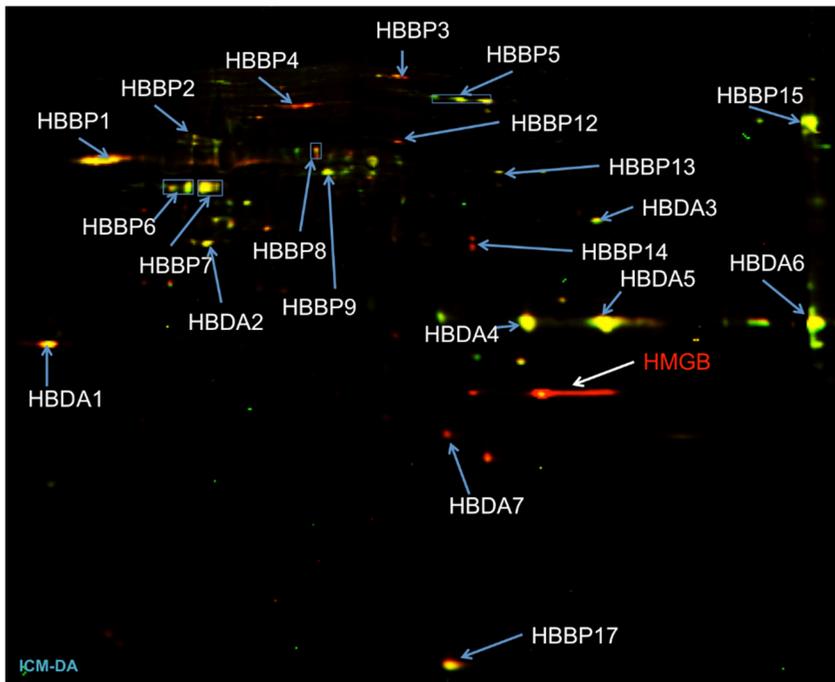
Supplementary Figure S2.2.10. Silver staining of 2D gel (pH 4–7 range) labeled by PPAR-BP.



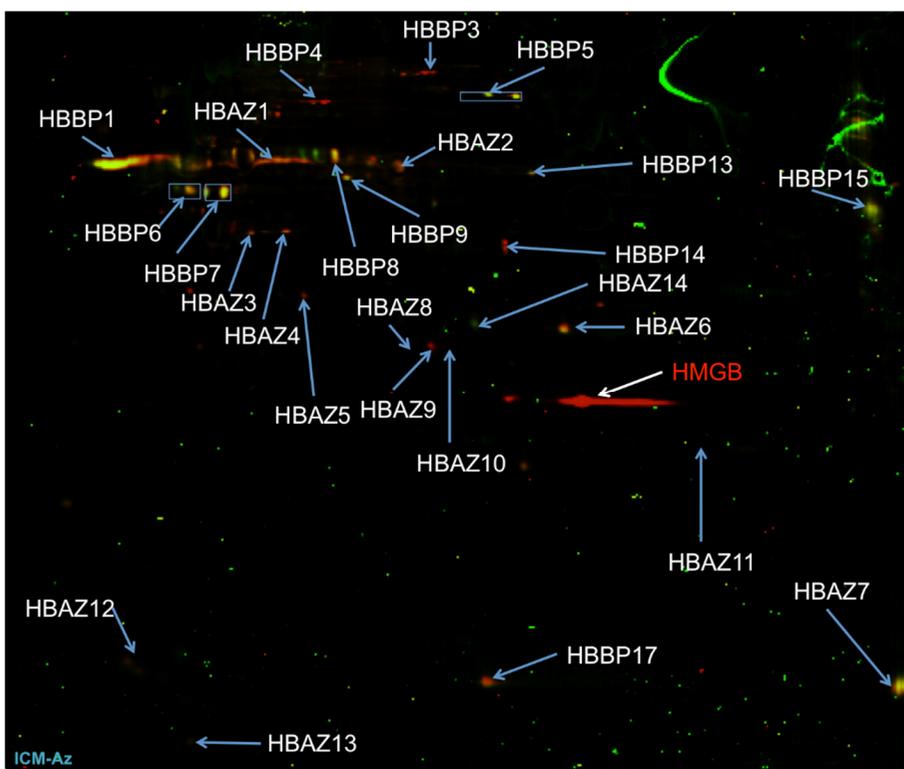
Supplementary Figure S2.2.11. Silver staining of 2D gel (pH 3–10 range) labeled by Tubulin-BP.



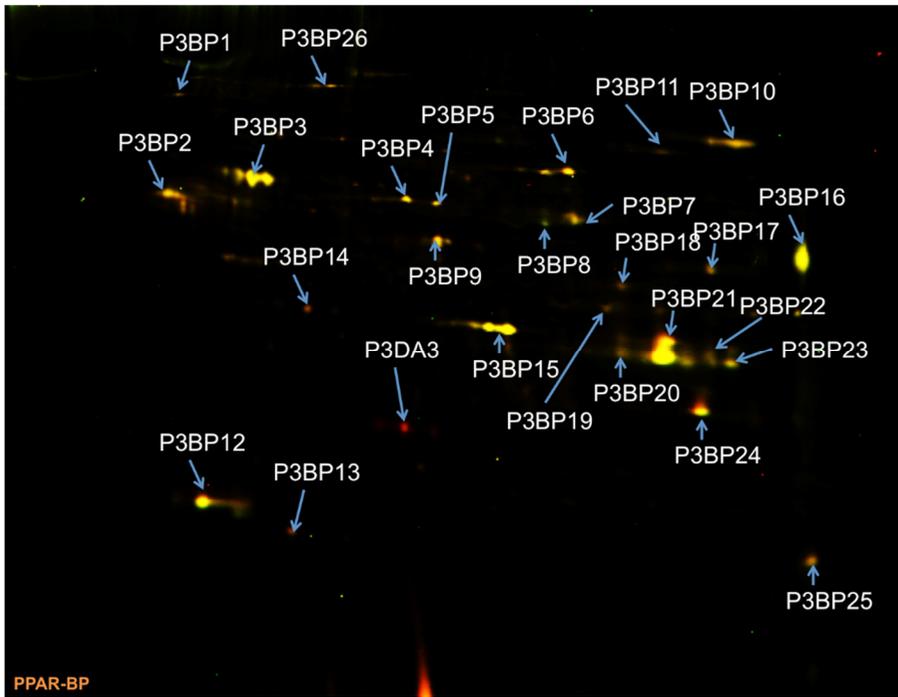
Supplementary Figure S2.2.12. Protein spots that were subjected for MS analysis and those file names in ICM target ID using ICM-BP (Figure 2.2.4C, left panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



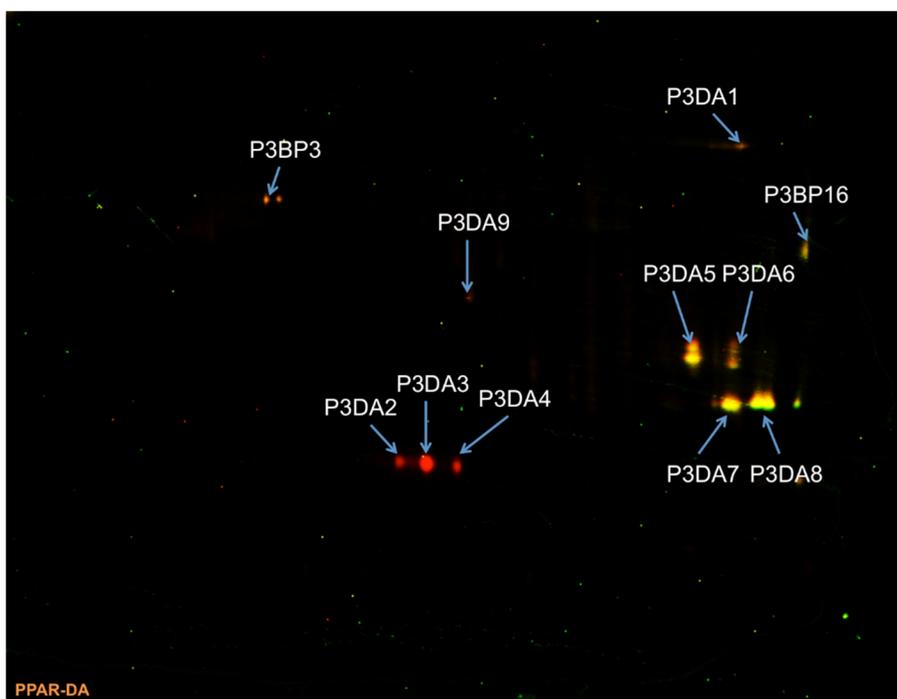
Supplementary Figure S2.2.13. Protein spots that were subjected for MS analysis and those file names in ICM target ID using ICM-DA (Figure 2.2.4C, center panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



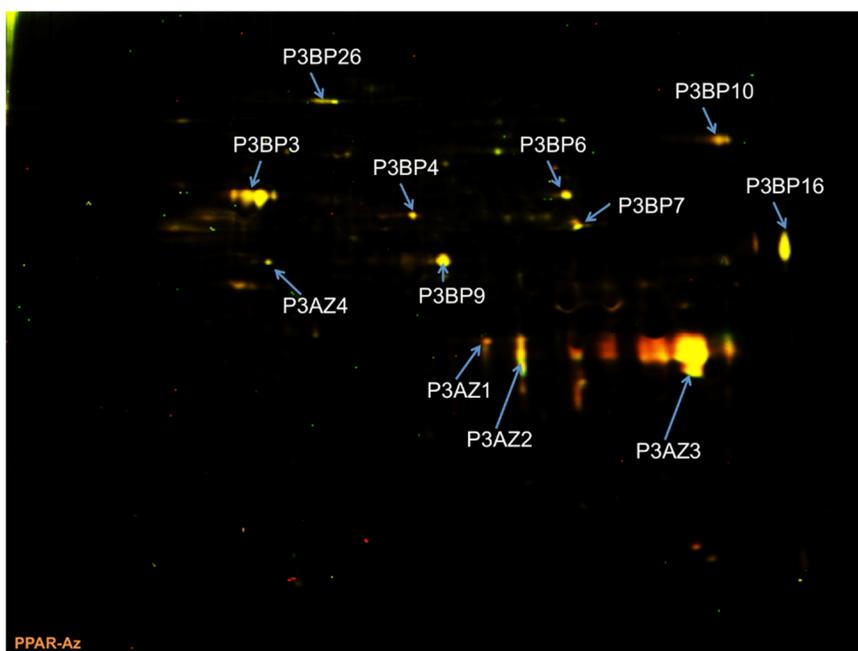
Supplementary Figure S2.2.14. Protein spots that were subjected for MS analysis and those file names in ICM target ID using ICM-Az (Figure 2.2.4C, right panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



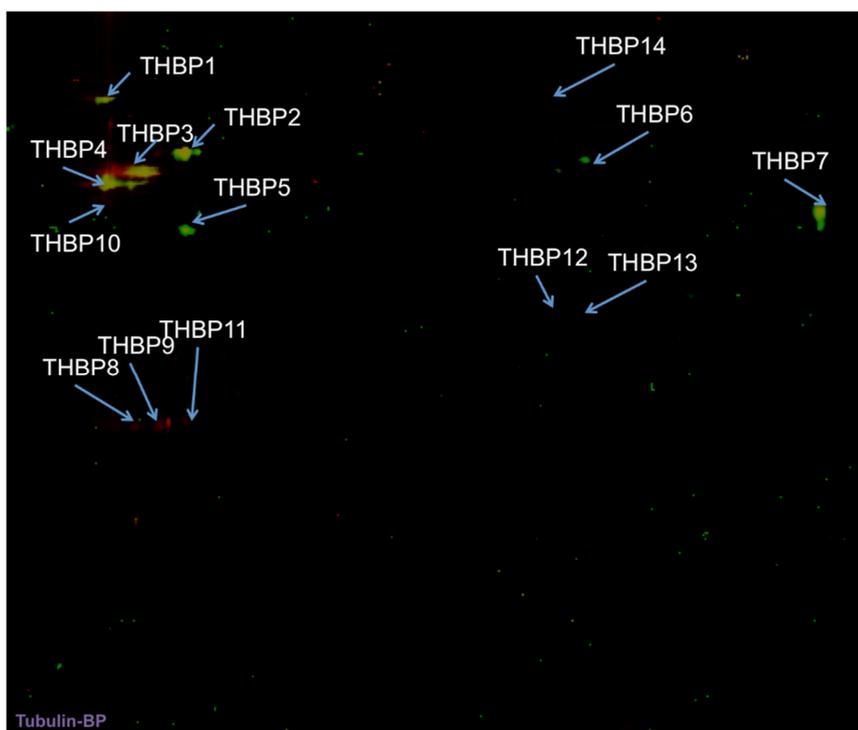
Supplementary Figure S2.2.15. Protein spots that were subjected for MS analysis and those file names in anti-diabetic agent **PPAR-e** target ID using PPAR-BP (Figure 2.2.5C, left panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



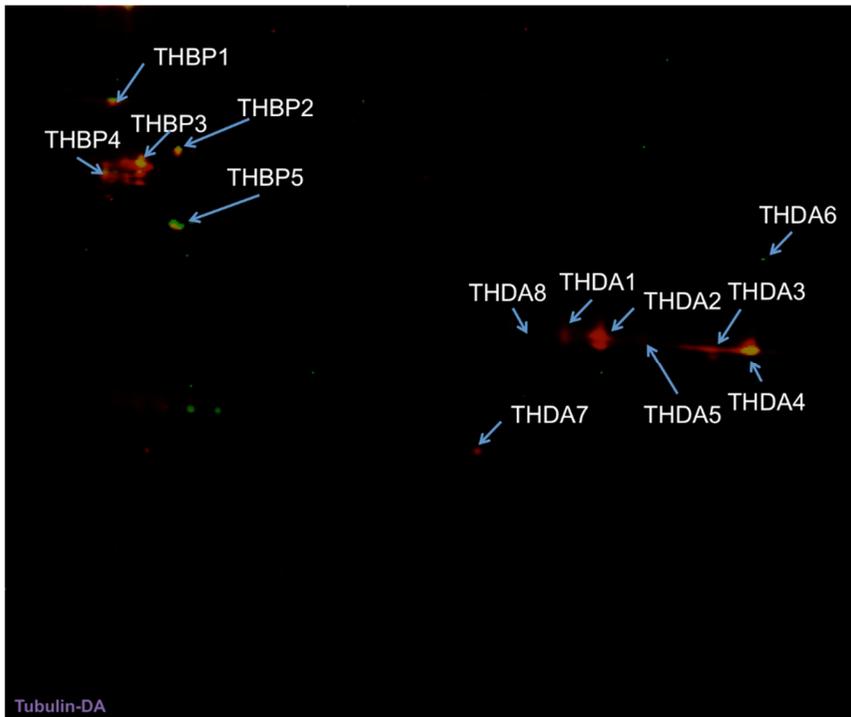
Supplementary Figure S2.2.16. Protein spots that were subjected for MS analysis and those file names in anti-diabetic agent **PPAR-e** target ID using PPAR-DA (Figure 2.2.5C, center panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



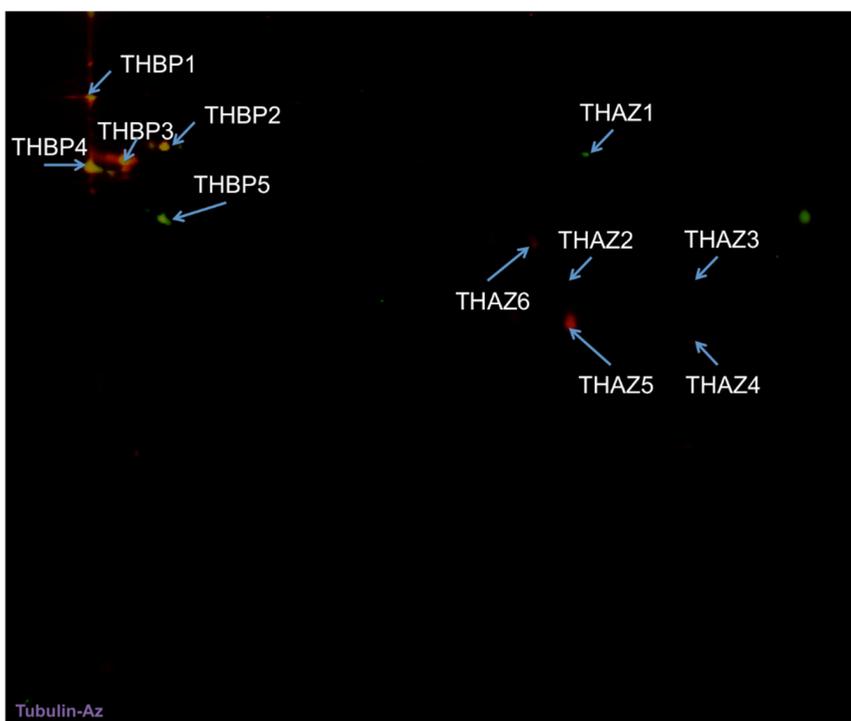
Supplementary Figure S2.2.17. Protein spots that were subjected for MS analysis and those file names in anti-diabetic agent **PPAR-e** target ID using PPAR-Az (Figure 2.2.5C, right panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



Supplementary Figure S2.2.18. Protein spots that were subjected for MS analysis and those file names in anti-cancer agent **Tubulin-i** target ID using Tubulin-BP (Figure 2.2.6C, left panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



Supplementary Figure S2.2.19. Protein spots that were subjected for MS analysis and those file names in anti-cancer agent **Tubulin-i** target ID using Tubulin-DA (Figure 2.2.6C, center panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.



Supplementary Figure S2.2.20. Protein spots that were subjected for MS analysis and those file names in anti-cancer agent **Tubulin-i** target ID using Tubulin-Az (Figure 2.2.6C, right panel). Each spot was cut and named for MS analysis. Protein MS data from each spot can be found in excel file by the name of its spot.

Photoaffinity linker	Protein name
1 specific binding proteins	vimentin DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 polyadenylate-binding protein heterogeneous nuclear ribonucleoprotein M HSP90 YTH domain family protein RCC2 protein alpha-hemoglobin HSP70 mitochondrial ATP synthase Pspc1NONO titin poly(A) binding protein CAP, adenylate cyclase-associated protein 1 eukaryotic translation initiation factor 4A enoyl-Coenzyme A hydratase lamin A/C transcript variant 1 NADH dehydrogenase (ubiquinone) flavoprotein 3 KIAA1150 protein splicing factor 3a TNF receptor-associated protein 1 elongation factor 1 alpha Chain C, Structure Of The Ran-Gppnhp-Ranbd1 Complex
2 specific binding proteins	L-lactate dehydrogenase Glyceraldehyde-3-phosphate dehydrogenase mitochondrial malate dehydrogenase 2 heterogeneous nuclear ribonucleoprotein A2/B1 RPS2 protein HNRPA1 protein Porin 31HM ribosomal protein S2 macropain triosephosphate isomerase Voltage-dependent anion-selective channel protein 3 nitric oxide synthase interacting protein PKM2 protein mitochondrial ribosomal protein L4 ankyrin repeat domain 23

3 specific binding proteins	FLJ00410 Perilipin-3 PAWR RAN binding protein 1
binding to linker 1, 2, 3	Tubulin actin
binding to linker 1,2	beta-globin
binding to linker 1,3	HSP60
Binding to linker 2,3	dermcidin

Supplementary Table S2.2.1. The lists of proteins binding to specific PLs. Most of PLs proteins specifically bind to certain PLs. Out of 54 PL binding proteins, only 5 proteins binds to more than 2 PLs.

Target ID probe	Protein name
ICM-BP	HSP90
ICM-Az	Actin HSP60
PPAR-BP	beta-globin elongation factor 1 alpha
PPAR-DA	L-lactate dehydrogenase Glyceraldehyde-3-phosphate dehydrogenase mitochondrial malate dehydrogenase 2 macropain triosephosphate isomerase
Tubulin-BP	HSP90 actin
Tubulin-DA	L-lactate dehydrogenase Glyceraldehyde-3-phosphate dehydrogenase mitochondrial malate dehydrogenase 2 heterogeneous nuclear ribonucleoprotein A2/B1 Tubulin Porin 31HM ribosomal protein S2 Voltage-dependent anion-selective channel protein

Supplementary Table S2.2.2. The list of proteins identified in additional red spots specifically labeled both by target ID probes and photoaffinity linkers. From our proteome-labeling profile of photoaffinity linkers, we could narrow down the number of target protein candidates during target ID process.

2.2.7. References and Notes

- † Portions of this chapter have been previously reported, see: Park, J.; Koh, M.; Koo, J. Y.; Lee, S.; Park, S.B. *ACS Chem. Biol.* **2016**, *11*, 44–52.
- (1) Swinney, D. C.; Anthony, J. *Nat. Rev. Drug Discov.* **2011**, *10*, 507–519.
- (2) Eder, J.; Sedrani, R.; Wiesmann, C. *Nat. Rev. Drug Discov.* **2014**, *13*, 577–587.
- (3) Welch, E. M.; Barton, E. R.; Zhuo, J.; Tomizawa, Y.; Friesen, W. J.; Trifillis, P.; Paushkin, S.; Patel, M.; Trotta, C. R.; Hwang, S.; Wilde, R. G.; Karp, G.; Takasugi, J.; Chen, G.; Jones, S.; Ren, H.; Moon, Y. C.; Corson, D. L.; Turpoff, A. A.; Campbell, J. A.; Conn, M. M.; Khan, A.; Almstead, N. G.; Hedrick, J.; Mollin, A.; Risher, N.; Weetall, M.; Yeh, S.; Branstrom, A. A.; Colacino, J. M.; Babiak, J.; Ju, W. D.; Hirawat, S.; Northcutt, V. J.; Miller, L. L.; Spatrack, P.; He, F.; Kawana, M.; Feng, H.; Jacobson, A.; Peltz, S. W.; Sweeney, H. L. *Nature* **2007**, *447*, 87–91.
- (4) Rottmann, M.; McNamara, C.; Yeung, B. K. S.; Lee, M. C. S.; Zou, B.; Russell, B.; Seitz, P.; Plouffe, D. M.; Dharia, N. V.; Tan, J.; Cohen, S. B.; Spencer, K. R.; González-Páez, G. E.; Lakshminarayana, S. B.; Goh, A.; Suwanarusk, R.; Jegla, T.; Schmitt, E. K.; Beck, H.; Brun, R.; Nosten, F.; Renia, L.; Dartois, V.; Keller, T. H.; Fidock, D. A.; Winzeler, E. A.; and Diagana, T. T. *Science* **2010**, *329*, 1175–1180.
- (5) Schenone, M.; Dančík, V.; Wagner, B. K.; Clemons, P. A. *Nat. Chem.*

- Biol.* **2013**, *9*, 232–240.
- (6) Park, J.; Koh, M.; Park, S. B. *Mol. BioSyst.* **2013**, *9*, 544–550.
- (7) Sato, S.; Kwon, Y.; Kamisuki, S.; Srivastava, N.; Mao, Q.; Kawazoe, Y.; Uesugi, M. *J. Am. Chem. Soc.* **2007**, *129*, 873–880.
- (8) Lim, H.; Cai, D.; Archer, C. T.; Kodadek, T. *J. Am. Chem. Soc.* **2007**, *129*, 12936–12937.
- (9) Hughes, C. C.; Yang, Y.; Liu, W.; Dorrestein, P. C.; Clair, J. J. L.; Fenical, W. *J. Am. Chem. Soc.* **2009**, *131*, 12094–12096.
- (10) Lomenick, B.; Hao, R.; Jonai, N.; China, R. M.; Aghajana, M.; Warburton, S.; Wang, J.; Wu, R. P.; Gomez, F.; Loo, J. A.; Wohlschlegel, J. A.; Vondriska, T. M.; Pelletier, J.; Herschman, H. R.; Clardy, J.; Clarke, C. F.; and Huang, J. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21984–21989.
- (11) Li, L.; Zhang, Q.; Liu, A.; Li, X.; Zhou, H.; Liu, Y.; Yan, B. *J. Am. Chem. Soc.* **2011**, *133*, 6886–6889.
- (12) Wacker, S. A.; Kashyap, S.; Li, X.; Kapoor, T. M. *J. Am. Chem. Soc.* **2011**, *133*, 12386–12389.
- (13) Li, Z.; Hao, P.; Li, L.; Tan, C. Y. J.; Cheng, X.; Chen, G. Y. J.; Sze, S. K.; Shen, H.-M.; Yao, S. Q. *Angew. Chem. Int. Ed.* **2013**, *52*, 8551–8556.
- (14) Li, Z.; Wang, D.; Li, L.; Pan, S.; Na, Z.; Tan, C. Y. J.; Yao, S. Q. *J. Am. Chem. Soc.* **2014**, *136*, 9990–9998.
- (15) Li, C.; Dong, T.; Li, Q.; Lei, X. *Angew. Chem. Int. Ed.* **2014**, *53*, 12111–12115.

- (16) Burdine, L.; Kodadek, T. *Chem. Biol.* **2004**, *11*, 593–597.
- (17) Rix, U.; Superti-Furga, G. *Nat. Methods* **2009**, *5*, 616–624.
- (18) Koh, M.; Park, J.; Koo, J. Y.; Lim, D.; Cha, M. Y.; Jo, A.; Choi, J. H.; Park, S. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 5102–5106.
- (19) Yang, P. Y.; Liu, K.; Ngai, M. H.; Lear, M. J.; Wenk, M. R.; Yao, S. *Q. J. Am. Chem. Soc.* **2010**, *132*, 656–666.
- (20) Shi, H.; Zhang, C. J.; Chen, G. Y. J.; Yao, S. *Q. J. Am. Chem. Soc.* **2012**, *134*, 3001–3014.
- (21) Kotzyba-Hibert, F.; Kapfer, I.; Goeldner, M. *Angew. Chem. Int. Ed.* **1995**, *34*, 1296–1312.
- (22) Smith, E.; Collins, I. *Future Chem. Med.* **2015**, *7*, 159–183.
- (23) Cisar, J. S.; Cravatt, B. F. *J. Am. Chem. Soc.* **2012**, *134*, 10385–10388.
- (24) Kambe, T.; Correia, B. E.; Niphakis, M. J.; Cravatt, B. F. *J. Am. Chem. Soc.* **2014**, *136*, 10777–10782.
- (25) Li, G.; Liu, Y.; Liu, Y.; Chen, L.; Wu, S.; Liu, Y.; Li, X. *Angew. Chem. Int. Ed.* **2013**, *52*, 9544–9549.
- (26) Park, J.; Oh, S.; Park, S. B. *Angew. Chem. Int. Ed.* **2012**, *51*, 5447–5451.
- (27) Lee, S.; Nam, Y.; Koo, J. Y.; Lim, D.; Park, J.; Ock, J.; Kim, J.; Suk, K.; Park, S. B. *Nat. Chem. Biol.* **2014**, *10*, 1055–1060.
- (28) Sakurai, K.; Tawa, M.; Okada, A.; Yamada, R.; Sato, N.; Inahara, M.; Inoue, M. *Chem. Asian J.* **2012**, *7*, 1567–1571.
- (29) Nakamura, Y.; Miyatake, R.; Ueda, M. *Angew. Chem. Int. Ed.* **2008**, *47*, 7289–7292.

- (30) Hatanaka, Y. *Chem. Pharm. Bull.* **2015**, *63*, 1–12.
- (31) Li, G.; Liu, Y.; Yu, X.; Li, X. *Bioconjugate Chem.* **2014**, *25*, 1172–1180.
- (32) Shannon, D. A.; Banerjee, R.; Webster, E. R.; Bak, D. W.; Wang, C.; Weerapana, E. *J. Am. Chem. Soc.* **2014**, *136*, 3330–3333.
- (33) Kim, J.; Kim, H.; Park, S. B. *J. Am. Chem. Soc.* **2014**, *136*, 14629–14638.
- (34) Oh, S.; Park, S. B. *Chem. Commun.* **2011**, *47*, 12754–12761.
- (35) Chuang, H. Y.; Chang, J. Y.; Lai, M. J.; Kuo, C. C.; Lee, H. Y.; Hsieh, H. P.; Chen, Y. J.; Chen, L. T.; Pan, W. Y.; Liou, J. P. *ChemMedChem* **2011**, *6*, 450–456.

2.3. Nonspecific Protein Labeling of Photoaffinity Linkers Correlates with Their Molecular Shapes in Living Cells

2.3.1. Abstract

Herein we report molecular shape-dependent nonspecific labeling of photoaffinity linkers (PLs) in the cellular proteome. Linear PLs have a higher tendency to engage in nonspecific binding than branched PLs. Exploiting this property, we discovered a smaller branched diazirine-based PL for the best photoaffinity probe with minimal nonspecific binding characteristics from among 5 probes with different PLs.

2.3.2. Introduction

Photoaffinity-based approach has received attention over the last few decades for target identification and activity-based protein profiling.¹ In the conventional affinity-based protein pull-down approach, bioactive small molecules with excellent potency (subnanomolar range activity) are generally required for successful enrichment of the interacting proteins.² Even with highly potent bioactive compounds, experimental conditions such as salt concentration, type of detergents, and temperature often interfere with the protein–small molecule interaction. This may lead to unexpected experimental failure in target identification.³ In contrast, the photoaffinity-based target

identification approach secures the interaction between small-molecule probes and their binding proteins under any experimental condition via the formation of UV-mediated covalent bonds. This approach expands the range of target identification from only high-affinity molecules to low-affinity molecules (micromolar range activity).⁴ However, nonspecific protein labeling by photoaffinity linkers (PLs) has been a major concern in the photoaffinity-based approach.⁵⁻⁹

2.3.3. Results and Discussion

Although a list of reports have emphasized the importance of probe design strategy for photoaffinity-based target identification, the molecular shape of PLs has been overlooked in the design of photoaffinity probes. Initially, we hypothesized that the molecular shape of PLs determines their flexibility, resulting in different protein-labeling patterns caused by nonspecific interaction. Following this initial hypothesis, we attempted to find a PL with minimal nonspecific labeling. To investigate molecular shape-dependent protein-labeling, we designed and synthesized 5 different PLs (**1**, **2**, **3**, **4**, and **5**) containing commonly used photoactivatable moieties (benzophenone and diazirine) and an acetylene moiety for bioorthogonal fluorescence labeling via Click chemistry (Figure 2.3.1a). Benzophenone-based PLs **1** and **2** have different shapes (**1**: linear, **2**: branched), as do diazirine-based PLs **3** and **4**. In addition, we designed and synthesized a branched PL **5** with the same number of total carbon and nitrogen atoms as the linear PL **3**, but smaller than another branched PL **4**. *In silico* analysis of surface charges showed clear differences

between linear and branched PLs (Figure 2.3.1b), which might lead to different noncovalent interactions with various proteins. The difference in molecular shapes resulted in a different number of conformers. As shown in Figure 2.3.1c, we simulated and counted the number of PL conformers having less than 2 kcal/mol difference from its energy minimum. The numerical value of 2 kcal/mol was derived from our assumption that conformers of each PL within a range of 2 kcal/mol from energy minimum can freely exist in the cellular environment under the physiological condition. The calculation results showed that linear PLs (**1** and **3**) have more conformers than branched PLs (**2**, **4**, and **5**). The conformational flexibility of linear PLs might cause the nonspecific, indiscriminate interactions via induced fit with various proteins. Based on their molecular shapes, we hypothesized that linear PLs would show more nonspecific protein labeling than branched PLs.

To examine the actual protein-labeling patterns of each PL, we mixed the 5 different PLs with either cell lysates or living cells. The mixtures were then subjected to UV irradiation to fix their transient binding with proteins by generating covalent bonds with the bound proteins. Proteins crosslinked with each PL were labeled with Cy5-azide using copper-mediated bioorthogonal Click chemistry and visualized in a fluorescent scanner after 1D-gel electrophoresis. Each PL showed different protein-labeling patterns in cell lysates and living cells (Figure 2.3.1d). As expected from calculated data, the protein-labeling patterns, particularly in living cells, were highly dependent on the molecular shapes of PLs. In living cells, branched PLs resulted in less nonspecific protein labeling than linear PLs. We also observed that the smaller

branched PL **5** showed significantly reduced nonspecific labeling in comparison with the larger branched PL **4** (Figure 2.3.1d, right panel, **4** and **5**), which was surprising given that these PL structures differ by only 4 methylene units (**4**: C₂₁H₃₅N₅O₄, **5**: C₁₇H₂₇N₅O₄). This result is in accordance with the findings of a previous study reporting the importance of small PLs.^{6a} In cell lysates, nonspecific protein labeling by PL **2** was significantly lower than that by PL **1**, but the protein-labeling patterns of diazirine-based PLs (**3**, **4**, and **5**) were generally similar and showed a little difference in protein bands in accordance with changes of their molecular shapes. The overall differences in protein-labeling patterns by various diazirine-based PLs in cell lysates were not as marked as that in living cells, which might be due to the fact that the diazirine-based PL conformers are not as diverse as we expected. Considering the much higher protein concentrations and complex protein networks with various protein–protein interactions in living cells,¹⁰ protein complex surfaces inside living cells should be more diverse than those in cell lysate. We believe that the diverse protein surfaces in living cells delineate the difference in the protein-labeling patterns of diazirine-based PLs. This observation was confirmed again by 2D-gel electrophoresis (Figure S2.3.2–S2.3.3). Examination of the nonspecific protein-labeling patterns of the 5 different PLs showed that in living cells, branched PLs tended to bind fewer proteins than linear PLs, which is in accordance with our original hypothesis regarding the molecular shape and conformational flexibility of PLs (Figure 2.3.1e).

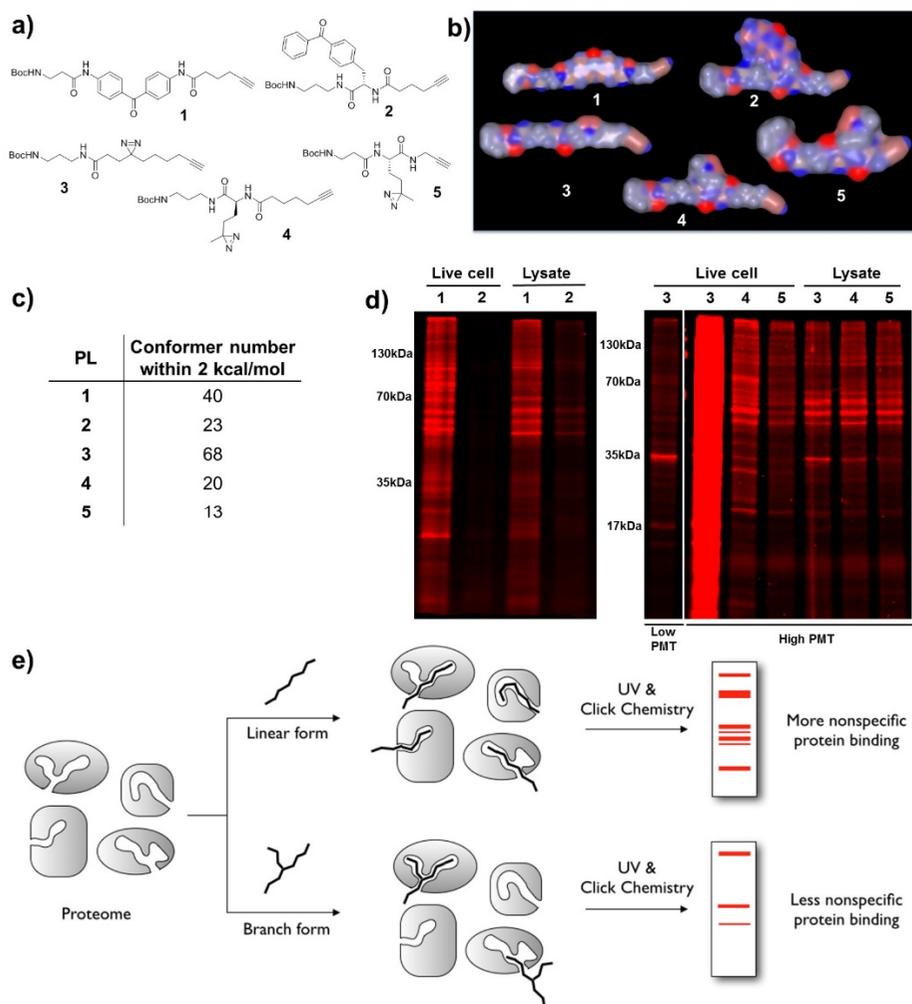


Figure 2.3.1. (a) Chemical structure of 5 photoaffinity linkers (PLs) with different molecular shapes. (b) Surface modeling of the 5 PLs. (c) The conformer number of PLs near the energy minimum. (d) The proteome labeled using PLs 1, 2, 3, 4, and 5 (10 μ M each) in living cells and cell lysates was subjected to 1D-gel electrophoresis for detailed analysis. A single SDS-PAGE gel with PLs 3, 4, and 5 was scanned with 2 different PMTs due to high nonspecific binding of PL 3. (e) Schematic illustration of the molecular shape-dependence of protein labeling. The flexibility of linear molecules increases the binding to various proteins. Branched molecules bind to fewer proteins than linear molecules due to their restricted conformational flexibility.

Based on this observation, we predicted that photoaffinity probes with branched PLs might show reduced nonspecific labeling and increased specific binding of the target proteins. To test this assumption, we designed a series of photoaffinity probes with various PLs using compound **6**, reported as a tubulin-targeting anticancer agent.¹¹ We first synthesized photoaffinity probes **7** and **8** by attaching benzophenone-based PLs **1** and **2** to the derivatives of **6** (Figure 2.3.2a). Then, we compared the protein-labeling patterns of probes **7** and **8** in either cell lysates or living cells in the absence or presence of agent **6**. After photo-crosslinking, the proteomes were labeled with Cy5-azide and analyzed using 1D-gel electrophoresis. Probe **7** (with a linear PL) showed more nonspecific protein labeling than probe **8** (with a branched PL) in both cell lysates and living cells (Figure 2.3.2b). The treatment concentration of each probe and agent **6** was chosen according to our previous report^{11,13} and a dose-dependent study (see Supporting Information). Interestingly, these protein-labeling patterns were consistent with those of PLs **1** and **2** themselves. However, tubulin was not selectively labeled by either probe **7** or **8** in cell lysates. In contrast, the overall protein-labeling patterns in living cells by probe **7** were not perturbed by an excess amount of soluble ligand **6** in the 1D gel. But, fewer proteins were labeled by **8** contained a branched PL than by **7**. Moreover, the target protein tubulin (Figure 2.3.2b: asterisk) was clearly labeled by **8**; the labeling of target proteins by probe **8** was inhibited in the presence of soluble ligand **6**, which was not observed for probe **7**. Therefore, we concluded that the protein-labeling patterns can be significantly affected by changing the molecular shapes of the PLs in the photoaffinity probes.

To improve the resolution of the protein analysis, we applied the dual-color system, running 2 different samples on a single and the largest 2D gel,¹¹ which ruled out the gel-to-gel variation. We used Cy5-azide or Cy3-azide for bioorthogonal reaction with proteins labeled with each photoaffinity probe in the absence or presence of **6**, respectively. Proteins labeled in yellow (Cy3+Cy5) or green (Cy3) are likely to be nonspecific binders. In contrast, proteins showing as red (Cy5) are probably potential target proteins of soluble ligand **6**. This substantially simplifies the target identification process. As observed in the 1D gel, the photoaffinity probe **7** showed more nonspecific binding than probe **8** in both cell lysates and living cells (Figure 2.3.2c) and intensive protein labeling by probe **7**, but not by **8**, was observed in cell lysates. In terms of specificity/selectivity, both photoaffinity probes (**7** and **8**) failed to label the target protein tubulin in cell lysates (Figure 2.3.2c and 2.3.2d), which demonstrates the importance of the cellular environment for the interaction between proteins and small molecules.¹² In living cells, the protein-labeling intensities by probes **7** and **8** were not very different. However, the number of protein spots labeled by **7** was larger, as observed in 1D gel (Figure 2.3.2b and 2.3.2c). The enlarged sections of 2D-gel images showed that not only the number of protein spots but also the labeling patterns depended on the molecular shapes of PLs (Figure 2.3.2d). Protein *a* (vimentin) and protein *b* (tubulin) were labeled in yellow by **7**, but in red by **8**. To identify the protein, we excised the protein spot from 2D gel and trypsinized for mass spectrometry (MS) analysis (see Supporting Information). The immunostaining results showed that **6** did not functionally modulate the vimentin filament (Figure

S2.3.10). Although the probe **8** could label tubulin, its selectivity for target protein was good enough to be used as a tubulin probe.

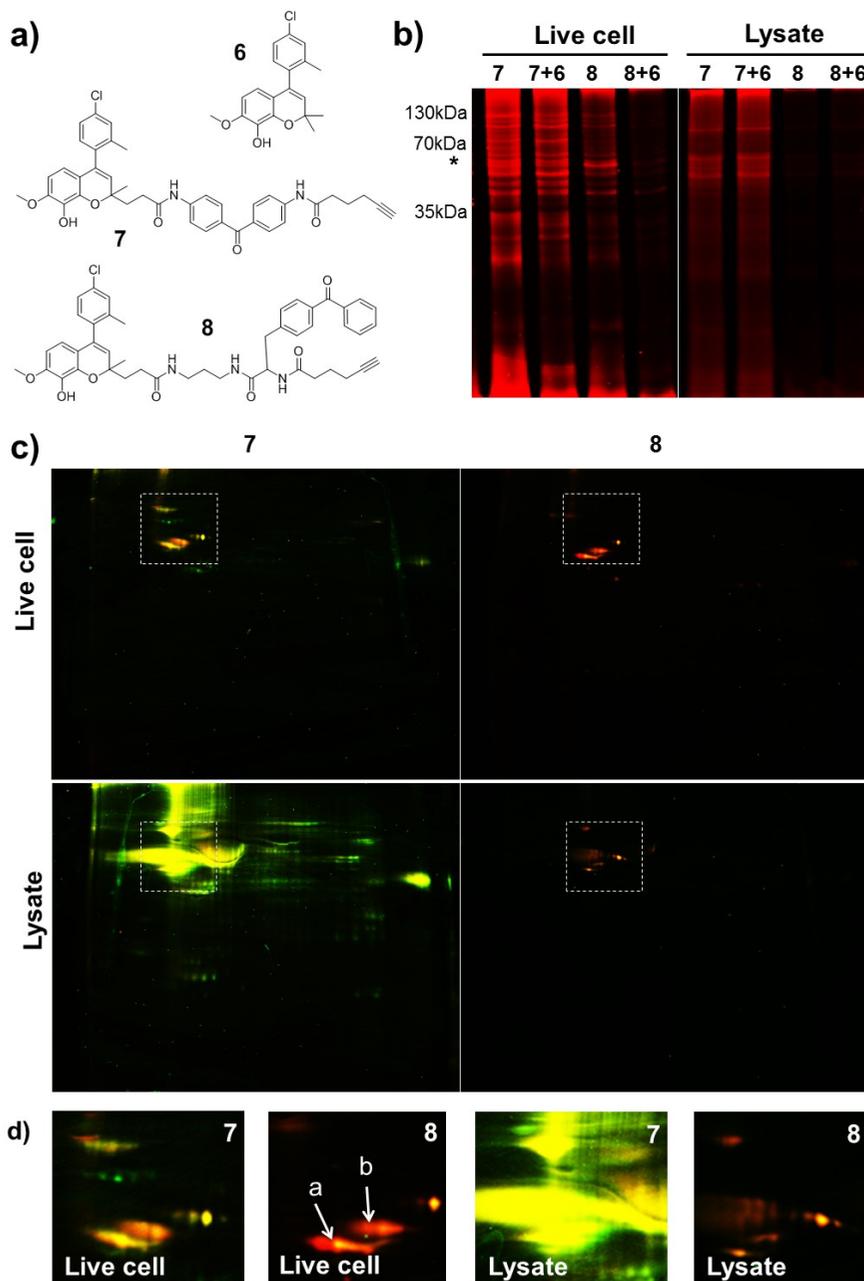


Figure 2.3.2. (a) Structure of tubulin inhibitor **6** and its photoaffinity probes **7** and **8** with

differently shaped benzophenone-based PLs. (b) The protein-labeling pattern for probes **7** and **8** on 1D-gel electrophoresis (**7**, **8**: 10 μ M, **6**: 40 μ M, *: target protein tubulin). (c) The protein-labeling pattern for probes **7** and **8** on 2D-gel electrophoresis (**7**, **8**: 10 μ M, **6**: 40 μ M). Red fluorescence: proteins labeled by probe **7** or **8**. Green fluorescence: proteins labeled by probe **7** or **8** in the presence of agent **6**. (d) Part of each 2D gel (white dashed rectangle in Fig. 2.3.2c) was magnified. Target and off-target protein labeling is dependent on the molecular shape of PLs (*a*: vimentin, *b*: tubulin).

To find the PL with the best selectivity, we moved our attention to diazirine-based PLs. Recent studies have reported the superior properties of diazirine-based PLs in comparison with benzophenone-based PLs.^{6a,14} Thus, we decided to reduce the nonspecific labeling of photoaffinity probes by using diazirine-based PLs (**3**, **4**, and **5**). We synthesized 3 different photoaffinity probes **9**, **10**, and **11** (Figure 2.3.3a). In cell lysates, an excess amount of soluble ligand **6** did not affect protein labeling by probes **9**, **10**, and **11**. Protein labeling intensities of **9** and **11** were similar to each other, which was in accordance with the results observed for protein-labeling patterns of PLs **3** and **5** (Figure 2.3.3b, right panel). Protein labeling of probe **10** is a little higher than that of probes **9** and **11**. We suspect the ligand part of probe **10** increased the overall protein labeling. In living cells, the overall protein-labeling patterns created by probes **9**, **10**, and **11** were consistent with the patterns for PLs **3**, **4**, and **5** (Figure 2.3.3b, left panel). Photoaffinity probes with diazirine-based PLs (**9**, **10**, and **11**) were more specific than probes with benzophenone-based PLs (**7** and **8**). The target protein tubulin was clearly labeled by probes **9**, **10**, and **11**, which supports the superiority of diazirine-based probes over benzophenone-based ones.^{6a,14} The photoaffinity probe with a diazirine-based linear PL (**9**) labeled tubulin more

selectively than the probe with a benzophenone-based linear PL (**7**). Nonspecific protein labeling by branched probes (**8** and **10**) was also markedly reduced than that by linear probes (**7** and **9**). Most importantly, the smaller branched diazirine-based probe (**11**) eliminated more nonspecific binding than all other photoaffinity probes.

To analyze the protein-labeling patterns with higher resolution, we performed dual-color 2D-gel electrophoresis using **9**, **10**, and **11**. In cell lysates, none of the protein-labeling patterns for all these probes was affected by the presence of **6**. Most proteins were labeled in yellow, which agreed with the results of 1D-gel experiments (Figure 2.3.3b and 2.3.3c). In contrast, we observed clear differences among protein-labeling patterns by **9**, **10**, and **11** upon treatment with a large excess of soluble ligand **6** in living cells; several proteins were labeled in red (Figure 2.3.3c). The number of protein spots in red was reduced as PLs were changed from linear (**9**) to branched (**10**) and to smaller branched PLs (**11**). Among proteins labeled by probe **9**, one of the proteins located on the right-hand side of the 2D gel was lactate dehydrogenase, a specific binder of a linear diazirine-based PL, which is not specific to agent **6** as we previously reported.¹³ Other proteins labeled by **9**, **10**, and **11** were located in the upper left region of the 2D gel (dotted rectangle). In the magnified images (Figure 2.3.3d), the color-labeling patterns of three protein spots were determined by the molecular shape of PLs in the photoaffinity probes. Probe **9** labeled proteins *a*, *b*, and *c* in red. Probe **10** labeled protein *a* and *b*, and probe **11** labeled protein *b* in red. Protein *a*, *b*, and *c* were excised from 2D gel and trypsinized for MS analysis (see Supporting Information). We have previously

demonstrated that vimentin (protein *a*, Figure S2.3.10) and HSP60 (protein *c*)¹¹ are not the target proteins of agent **6**. Only tubulin (protein *b*) is the target protein of agent **6**, confirmed by functional biochemical assays.¹¹ The photoaffinity probe **11** with a smaller branched PL selectively labeled the target protein tubulin. Probes **9** and **10** labeled tubulin, but they also bound other non-target proteins; such behavior increases the amount of time and efforts put in target validation. These results confirmed that photoaffinity probes with branched PLs are generally better than those with linear PLs. Moreover, a smaller branched PL in this study shows lower nonspecific labeling and higher selectivity toward target protein, tubulin.

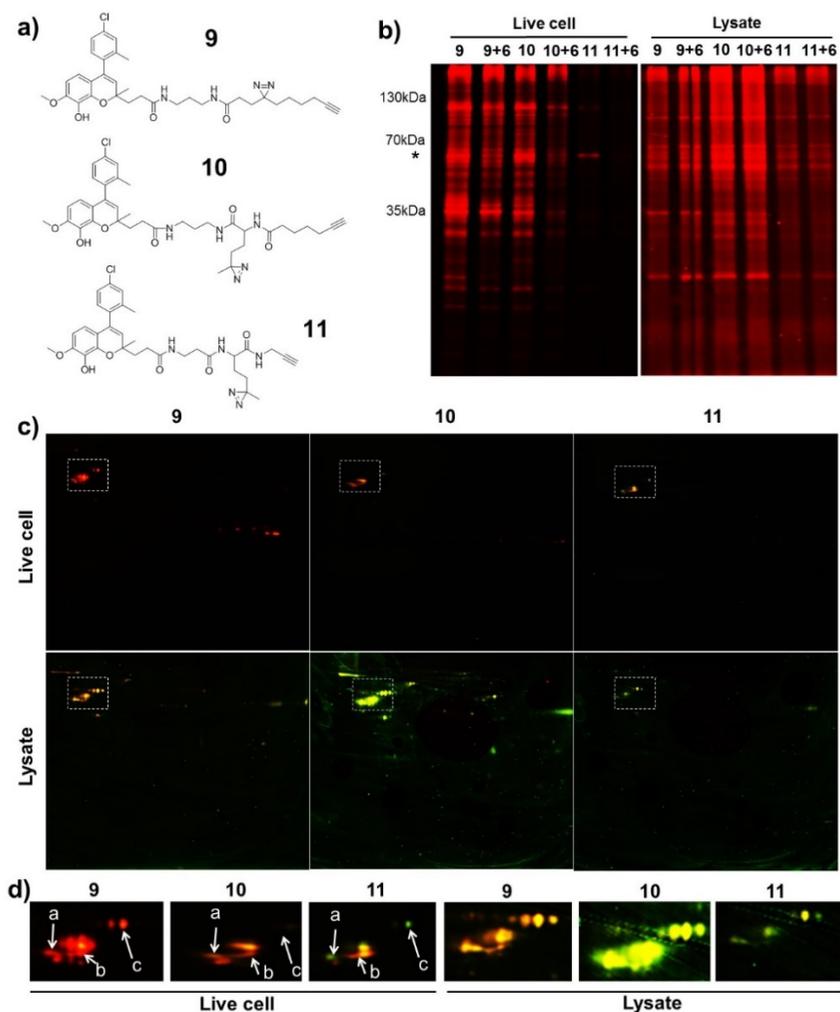


Figure 2.3.3. (a) Structure of photoaffinity probes **9**, **10**, and **11** with differently shaped diazirine-based PLs. (b) Protein-labeling patterns for probes **9**, **10**, and **11** on 1D-gel electrophoresis (**9**, **10**, **11**: 10 μ M, **6**: 10 μ M for living cells and 40 μ M for cell lysate, *: target protein tubulin). Probe **11** selectively labels tubulin in comparison with probes **9** and **10**. (c) Protein-labeling patterns for probes **9**, **10**, and **11** on 2D-gel electrophoresis (**9**, **10**, **11**: 10 μ M, **6**: 10 μ M for living cells and 40 μ M for cell lysate). Red fluorescence: protein labeled by probe **9**, **10**, or **11**. Green fluorescence: protein labeled by probe **9**, **10**, or **11** in the presence of agent **6**. (d) Part of each 2D gel (white dashed rectangular in Fig. 2.3.3c) was magnified. Target and off-target protein labeling is dependent on the molecular shape of PLs (*a*: vimentin, *b*: tubulin, *c*: HSP60)

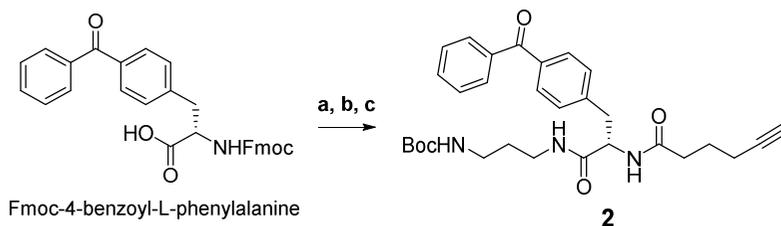
2.3.4. Conclusion

In conclusion, we systematically studied the molecular shape-dependent protein labeling of PLs. Direct comparison of 5 different PLs with various molecular shapes and photoactivatable moieties showed significantly reduced nonspecific protein labeling by branched PLs in comparison with linear PLs in living cells. This result is probably due to the high conformational flexibility of linear PLs. The application of these findings in the design of photoaffinity probes clearly confirmed the superiority of branched PLs in specific labeling procedures. Exploiting the molecular shape-dependent properties of PLs, we synthesized a tubulin-selective photoaffinity probe **11** and demonstrated that the photoaffinity-based approach with a well-designed probe can identify the target proteins in living cells effectively. Currently, we plan to apply our result to a series of bioactive compounds having known multiple target proteins to monitor the labeling efficiency of compactly branched PL toward each target protein.

2.3.5. Experimental Section

Preparation and characterization of all new compounds (**2**, **4**, **5**, and **8–15**) were described. Preparation and characterization of **1**, **3**, **7**, and **9** were previously reported.^{11,13}

Supplementary Scheme S2.3.1. Synthesis of Benzophenone Linker 2



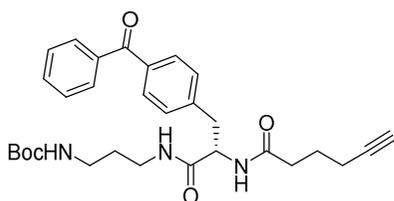
Reagents and conditions: (a) *N*-Boc-1,3-propanediamine, HATU, DIPEA, DMF, 0 °C, 1 h; (b) Piperidine, DMF, r.t., 30 min; (c) 5-Hexynoic acid, HATU, DIPEA, DMF, r.t., 1 h, overall yield: 31%.

(Step **a**) To a 3 mL anhydrous DMF solution of Fmoc-4-benzoyl-L-phenylalanine (0.41 mmol, 1 equiv.) and HATU (1.2 equiv.), *N*-Boc-1,3-propanediamine (1.1 equiv.) and DIPEA (1.5 equiv.) were added at 0 °C. The reaction mixture was stirred at 0 °C for 1 h. Then, the reaction solution was diluted with ethyl acetate and water. The aqueous layer was extracted three times with ethyl acetate. The organic layer was evaporated under reduced pressure to provide the crude amine product. The combined organic layer was dried over anhydrous $\text{MgSO}_4(\text{s})$ and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and was directly for next reaction without further purification.

(Step **b**) To a solution of the resulting crude product in DCM (5 mL), excess amount of piperidine (1 mL) was added. The solution was stirred at room temperature for 1 h. After completion of the reaction, the reaction mixture was concentrated *in vacuo* and purified with silica-gel flash column chromatography (DCM:methanol = 7:1) to provide the desired Fmoc-deprotected product. ($R_f=0.44$ at DCM:methanol = 8:1)

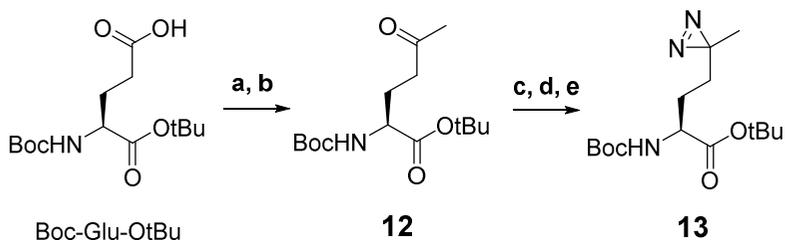
(Step **c**) To a resulting amine product of step **b** and 5-hexynoic acid (2 equiv.) in DMF (5 mL), HATU (1.5 equiv.) and diisopropylethylamine (3 equiv.) were added. The reaction mixture was stirred for 1 h at room temperature. Then, the reaction solution was diluted with ethyl acetate and water. The aqueous layer was extracted three times with ethyl acetate. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **2**. (3-step yield: 31%)

■ *tert*-Butyl (S)-3-(3-(4-benzoylphenyl)-2-(hex-5-ynamido)propanamido)propyl)carbamate (**2**)



$R_f=0.68$ (DCM:methanol = 8:1, v/v); ¹H NMR (500 MHz, CDCl₃) δ 7.69–7.75 (m, 4H), 7.56 (t, $J = 7.5$ Hz, 1H), 7.45 (t, $J = 7.8$ Hz, 1H), 7.31 (d, $J = 8.5$ Hz, 1H), 7.01 (brs, 1H), 6.59 (d, $J = 7.5$ Hz, 1H), 4.91 (brs, 1H), 4.75 (d, $J = 7.3$ Hz, 1H), 3.09–3.28 (m, 4H), 2.91–2.99 (m, 2H), 2.32 (t, $J = 7.3$ Hz, 1H), 2.12–2.18 (m, 2H), 1.93 (t, $J = 2.5$ Hz, 1H), 1.74–1.81 (m, 2H), 1.51 (brs, 2H), 1.38 (s, 9H); ¹³C NMR (125 MHz, CDCl₃) δ 196.5, 172.5, 171.0, 156.8, 142.0, 137.7, 136.4, 132.6, 130.6, 130.2, 129.5, 128.5, 83.5, 79.6, 69.6, 54.4, 38.8, 37.2, 36.2, 35.0, 30.1, 28.6, 24.2, 18.0; LRMS(ESI⁺) m/z calcd for C₃₀H₃₇N₃O₅ [M+H]⁺ 520.28; Found 520.3.

Supplementary Scheme S2.3.2. Synthesis of Diazirine Linker Precursor 13



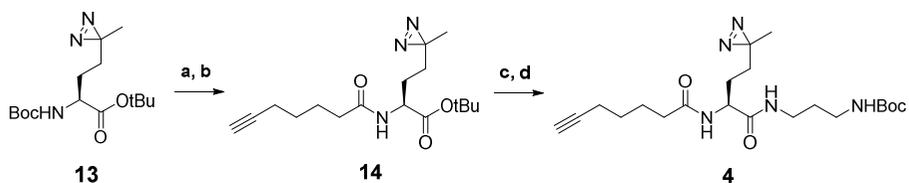
Reagents and conditions: (a) *N,O*-Dimethylhydroxylamine hydrochloride, NMM, EDC·HCl, DCM, $-15\text{ }^{\circ}\text{C} \rightarrow \text{r.t.}$, 12 h; (b) Methylmagnesium bromide solution, THF, $-78\text{ }^{\circ}\text{C} \rightarrow 0\text{ }^{\circ}\text{C}$, 12 h, 74%; (c) Liquid ammonia, $-78\text{ }^{\circ}\text{C} \rightarrow \text{r.t.}$, 12 h; (d) Hydroxylamine-*O*-sulfonic acid, methanol, r.t., 12 h; (e) I_2 , TEA, $0\text{ }^{\circ}\text{C}$, 2 h, 49%.

(Step **a**) To a solution of Boc-Glu-OtBu (2.0 g, 1 equiv.) in anhydrous DCM (20 mL) at $-15\text{ }^{\circ}\text{C}$, *N,O*-dimethylhydroxylamine hydrochloride (1.05 equiv.) and *N*-methylmorpholine (1.05 equiv.) were added. After stirring for 5 min, EDC hydrochloride (1.05 equiv.) in DCM (10 mL) was added over 15 min. Then the reaction mixture was stirred for 12 h at room temperature. Then the reaction mixture was diluted with DCM and ddH₂O. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and the resulting clear oil was directly for the subsequent substitution reaction without further purification.

(Step **b**) To a solution of a resulting crude mixture (from step **a**) in anhydrous THF (50 mL), 3.0 M solution of methylmagnesium bromide in diethyl ether (3 equiv.) was added at $-78\text{ }^{\circ}\text{C}$. The resulting mixture was then stirred for 3 h with the temperature warming up to $0\text{ }^{\circ}\text{C}$. Subsequently, the reaction was quenched

by addition of water (30 mL) followed by extraction with three times with DCM. The organic layer was washed with brine. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **12** (yield: 74%). NMR spectra of compound **12** was in agreement with that reported in previous literature.¹⁴ (Step **c**, **d**, **e**) The synthesis of **13** from **12** was performed using previously reported procedure.¹⁴

Supplementary Scheme S2.3.3. Synthesis of diazirine linker **4**



Reagents and conditions: (a) HCl, dioxane, r.t., 15 min; (b) 6-Heptynoic acid, EDC·hydrochloride, HOBT, TEA, r.t., 12 h, 42%; (c) TFA, DCM, r.t., 14 h; (d) *N*-Boc-1,3-propanediamine, EDC·hydrochloride, HOBT, TEA, r.t., 12 h, 63%.

(Step **a**) To a solution of **13** (0.75 mmol, 1 equiv.) in anhydrous 1,4-dioxane (3 mL), 4 N HCl solution in 1,4-dioxane (3 mL) was added and the reaction mixture was stirred at room temperature for 15 min. Saturated aqueous bicarbonate solution was added to the solution and the mixture was extracted with DCM. The organic layer was evaporated under reduced pressure to provide crude amine product.

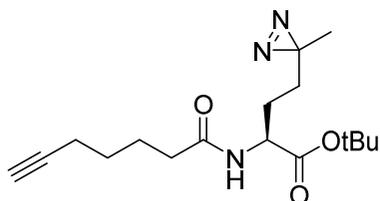
(Step **b**) To a resulting crude product of step **a** and 6-heptynoic acid (1.5 equiv.) in 1,4-dioxane (3 mL), EDC·hydrochloride (1.5 equiv.), hydroxybenzotriazole

(1.5 equiv.), and triethylamine (6 equiv.) were added. The reaction mixture was stirred for 12 h at room temperature. Then, the reaction solution was diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous $\text{MgSO}_4(\text{s})$ and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **14**. (2-step yield: 42%)

(Step **c**) Compound **14** (1 equiv.) was dissolved in 15% (v/v) trifluoroacetic acid in DCM (2 mL). The mixture was stirred at room temperature for 14 h and condensed under reduced pressure to provide substituted amine TFA salt. The resulting mixture was used directly for the subsequent amide coupling reaction without further purification.

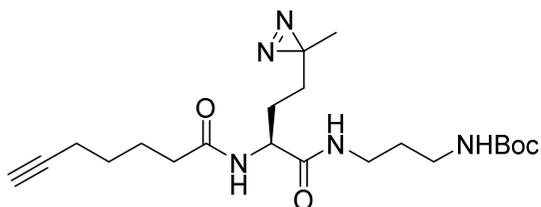
(Step **d**) To a resulting crude salt of step **c** and *N*-Boc-1,3-propanediamine (1.5 equiv.) in 1,4-dioxane (3 mL), EDC·hydrochloride (1.5 equiv.), hydroxybenzotriazole (1.5 equiv.), and triethylamine (6 equiv.) were added. The reaction mixture was stirred for 12 h at room temperature. Then, the reaction solution was diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous $\text{MgSO}_4(\text{s})$ and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **4**. (2-step yield: 63%)

- *tert*-Butyl (*S*)-2-(hept-6-ynamido)-4-(3-methyl-3*H*-diazirin-3-yl)butanoate (14)



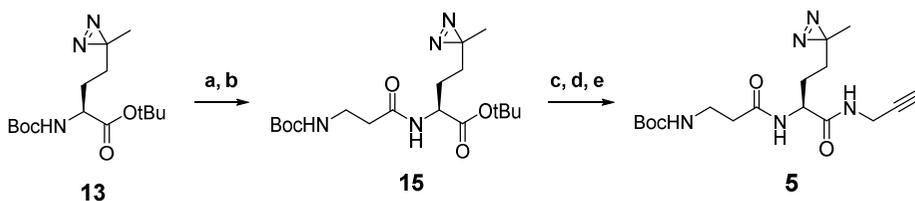
$R_f = 0.30$ (ethyl acetate:n-hexane = 1:3, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 6.00 (d, $J = 7.2$ Hz, 1H), 4.46 (q, $J = 6.5$ Hz, 1H), 2.18–2.25 (m, 4H), 1.95 (t, $J = 2.6$ Hz, 1H), 1.66–1.80 (m, 2H), 1.38–1.60 (m, 4H), 1.44 (s, 9H), 1.21–1.31 (m, 2H), 1.00 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 172.4, 171.3, 84.1, 82.7, 68.8, 52.0, 36.1, 30.5, 28.1, 28.0, 27.4, 25.4, 24.7, 19.7, 18.3.

- *tert*-Butyl (*S*)-3-(2-(hept-6-ynamido)-4-(3-methyl-3*H*-diazirin-3-yl)butanamido)propyl)carbamate (4)



$R_f = 0.75$ (DCM:methanol = 10:1, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.07 (brs, 1H), 6.35–6.38 (m, 1H), 4.91 (brs, 1H), 4.41 (q, $J = 6.5$ Hz, 1H), 3.26 (q, $J = 6.3$ Hz, 2H), 3.10–3.14 (m, 2H), 2.17–2.26 (m, 4H), 1.95 (t, $J = 2.4$ Hz, 1H), 1.68–1.78 (m, 2H), 1.39–1.64 (m, 6H), 1.43 (s, 9H), 1.23–1.34 (m, 2H), 1.00 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 172.9, 171.4, 156.8, 84.1, 79.6, 68.9, 52.5, 37.2, 36.1, 36.0, 30.6, 30.1, 28.5, 28.0, 27.5, 25.5, 24.7, 19.7, 18.3; LRMS(ESI $^+$) m/z calcd for $\text{C}_{21}\text{H}_{35}\text{N}_5\text{O}_4$ [M+H] $^+$ 422.28; Found 422.28.

Supplementary Scheme S2.3.4. Synthesis of diazirine linker **5**



Reagents and conditions: (a) HCl, dioxane, r.t., 15 min; (b) N-Boc- β -alanine, EDC hydrochloride, HOBT, TEA, 54%; (c) TFA, DCM, r.t., 12 h; (d) Di-*tert*-butyl dicarbonate, NaOH, water, THF, 20 h; (e) propargylamine, EDC-hydrochloride, HOBT, TEA, 1,4-dioxane, r.t., 12 h, 9%.

(Step **a**) To a solution of **13** (0.75 mmol, 1 equiv.) in anhydrous 1,4-dioxane (3 mL), 4 N HCl solution in 1,4-dioxane (3 mL) was added and the reaction mixture was stirred at room temperature for 15 min. Saturated aqueous bicarbonate solution was added to the solution and the mixture was extracted with DCM. The organic layer was evaporated under reduced pressure to provide the crude amine product.

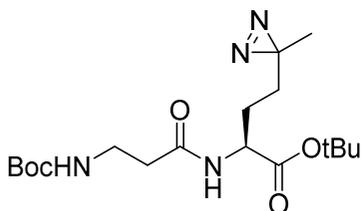
(Step **b**) To a resulting crude product of step **a** and N-Boc- β -alanine (1.5 equiv.) in 1,4-dioxane (3 mL), EDC-hydrochloride (1.5 equiv.), hydroxybenzotriazole (1.5 equiv.), and triethylamine (6 equiv.) were added. The reaction mixture was stirred for 12 h at room temperature. Then, the reaction solution was diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **15**. (2-step yield: 54%)

(Step **c**) Compound **15** (1 equiv.) was dissolved in 15% (v/v) trifluoroacetic acid in DCM (2 mL). The mixture was stirred at room temperature for 14 h and condensed under reduced pressure to provide substituted amine TFA salt. The resulting mixture was directly for the subsequent Boc-protection reaction without further purification.

(Step **d**) To a solution of the resulting crude salt of step **c** in water (3 mL), 2 N sodium hydroxide aqueous solution (3 mL) was added at room temperature followed by addition of di-*tert*-butyl dicarbonate (1.1 equiv.) in THF (6 mL). The reaction mixture was stirred for 20 h at room temperature and concentrated under reduced pressure. Excess equivalent of saturated aqueous ammonium chloride solution was added to the solution and the mixture was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* to provide crude Boc-protected product.

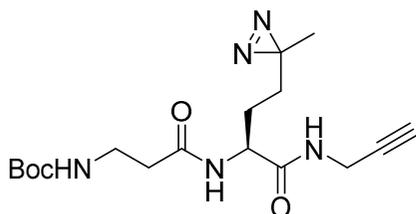
(Step **e**) To a resulting crude product of step **d** and propargylamine (1.5 equiv.) in 1,4-dioxane (3 mL), EDC·hydrochloride (1.5 equiv.), hydroxybenzotriazole (1.5 equiv.), and triethylamine (6 equiv.) were added. The reaction mixture was stirred for 12 h at room temperature. Then, the reaction solution was diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **5**. (3-step yield: 9%)

- *tert*-Butyl (*S*)-2-(3-((*tert*-butoxycarbonyl)amino)propanamido)-4-(3-methyl-3*H*-diazirin-3-yl)butanoate (**15**)



$R_f = 0.56$ (ethyl acetate:n-hexane = 1:1, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 6.14 (d, $J = 7.6$ Hz, 1H), 5.12 (brs, 1H), 4.40–4.46 (m, 1H), 3.39 (brs, 2H), 2.38–2.44 (m, 2H), 1.68–1.75 (m, 1H), 1.35–1.55 (m, 20H), 1.23–1.32 (m, 1H), 1.00 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 171.3, 171.1, 156.1, 82.8, 79.5, 52.1, 36.8, 36.3, 30.5, 28.5, 28.1, 27.2, 25.4, 19.7.

- *tert*-Butyl (*S*)-(3-((4-(3-methyl-3*H*-diazirin-3-yl)-1-oxo-1-(prop-2-yn-1-ylamino)butan-2-yl)amino)-3-oxopropyl)carbamate (**5**)



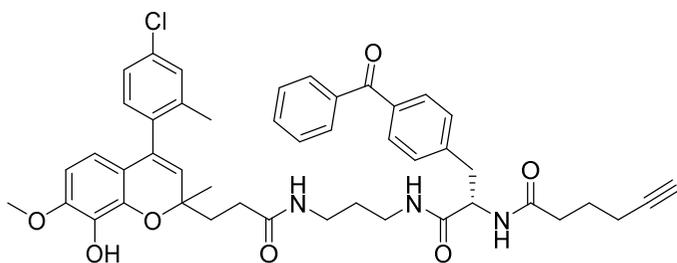
$R_f = 0.07$ (ethyl acetate:n-hexane = 1:1, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 6.62–6.66 (m, 1H), 6.41 (brs, 1H), 5.08 (brs, 1H), 4.41 (q, $J = 6.9$ Hz, 1H), 4.02 (q, $J = 2.7$ Hz, 2H), 3.40 (q, $J = 6.1$ Hz, 2H), 2.44 (q, $J = 5.3$ Hz, 1H), 2.24 (t, $J = 2.6$ Hz, 1H), 1.43–1.52 (m, 11H), 1.24–1.37 (m, 2H), 1.02 (s, 3H); LRMS(ESI $^+$) m/z calcd for $\text{C}_{17}\text{H}_{27}\text{N}_5\text{O}_4$ $[\text{M}+\text{H}]^+$ 366.21; Found 366.17.

[General synthetic procedure for compounds **8**, **9**, **10**, and **11**]

First, compound **2**, **3**, **4**, or **5** was dissolved in 15% (v/v) trifluoroacetic acid in DCM (2 mL). The mixture was stirred at room temperature for 12 h and condensed under reduced pressure to provide crude amine TFA salt. The resulting mixture was directly for the subsequent amide coupling reaction without further purification.

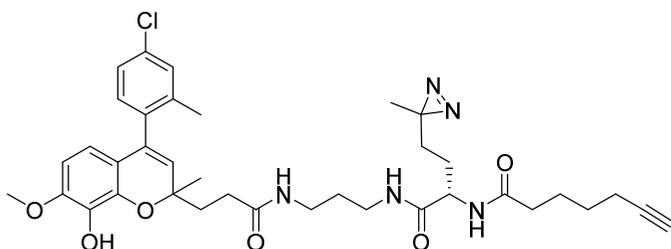
To a resulting crude amine salt (1.5 equiv.) of previous step and 3-(4-(4-chloro-2-methylphenyl)-8-hydroxy-7-methoxy-2-methyl-2*H*-chromen-2-yl)propanoic acid (1 equiv.)¹¹ in 1,4-dioxane (3 mL), EDC·hydrochloride (1.5 equiv.), hydroxybenzotriazole (1.5 equiv.), and triethylamine (6 equiv.) were added. The reaction mixture was stirred for 12 h at room temperature. Then, the reaction solution was diluted with DCM and water. The aqueous layer was extracted three times with DCM. The combined organic layer was dried over anhydrous MgSO₄(s) and filtered through a celite-packed glass filter. The filtrate was concentrated *in vacuo* and purified with silica-gel flash column chromatography to provide the desired product **8** (yield: 78%), **9** (yield: 59%), **10** (yield: 57%), or **11** (yield: 66%). In case of **10** and **11**, one single diastereomer was purified by normal-phase HPLC for NMR spectrum.

■ *N*-((2*S*)-3-(4-Benzoylphenyl)-1-((3-(3-(4-(4-chloro-2-methylphenyl)-8-hydroxy-7-methoxy-2-methyl-2*H*-chromen-2-yl)propanamido)propyl)amino)-1-oxopropan-2-yl)hex-5-ynamide (**8**)



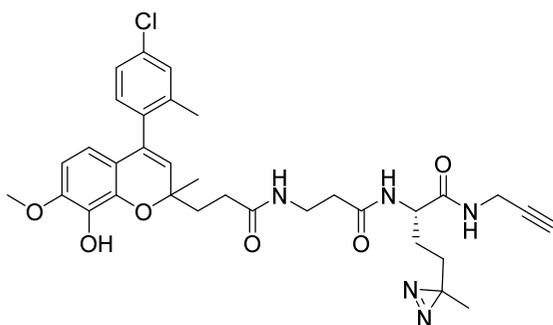
$R_f = 0.41$ (DCM:methanol = 10:1, v/v); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.68–7.75 (m, 4H), 7.56 (t, $J = 7.5$ Hz, 1H), 7.44 (td, $J = 7.6, 1.7$ Hz, 2H), 7.32 (d, $J = 8.5$ Hz, 2H), 7.12–7.23 (m, 2H), 6.98–7.07 (m, 2H), 6.38–6.66 (m, 2H), 6.30–6.32 (m, 1H), 6.05–6.08 (m, 1H), 5.29 (t, $J = 6.5$ Hz, 1H), 4.74–4.79 (m, 1H), 3.81 (s, 3H), 2.96–3.21 (m, 6H), 2.31–2.49 (m, 4H), 1.99–2.19 (m, 7H), 1.91–1.94 (m, 1H), 1.74–1.82 (m, 2H), 1.41–1.51 (m, 5H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 196.6, 173.9, 173.8, 172.6, 171.5, 171.3, 148.5, 141.9, 140.2, 140.0, 138.6, 138.4, 137.5, 136.4, 136.2, 134.8, 134.4, 134.3, 134.2, 133.5, 133.4, 132.6, 131.1, 130.9, 130.4, 130.1, 130.0, 129.9, 129.4, 128.4, 126.0, 126.0, 116.1, 116.0, 103.8, 103.8, 83.4, 79.0, 78.9, 69.5, 56.2, 54.5, 54.4, 38.6, 36.8, 36.6, 36.5, 36.2, 34.9, 31.8, 31.7, 29.8, 29.4, 26.3, 26.2, 26.0, 25.8, 24.2, 19.9, 19.6, 17.9 (peaks of two diastereomers); LRMS(ESI⁺) m/z calcd for $\text{C}_{46}\text{H}_{48}\text{ClN}_3\text{O}_7$ [M+H]⁺ 790.33; Found 790.3.

■ *N*-((2*S*)-1-((3-(3-(4-(4-chloro-2-methylphenyl)-8-hydroxy-7-methoxy-2-methyl-2*H*-chromen-2-yl)propanamido)propyl)amino)-4-(3-methyl-3*H*-diazirin-3-yl)-1-oxobutan-2-yl)hept-6-ynamide (**10**)



$R_f = 0.36$ (DCM:methanol = 10:1, v/v); $^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.26 (d, $J = 5.2$ Hz, 1H), 7.22 (d, $J = 8.4$ Hz, 1H), 7.08–7.13 (m, 1H), 6.40 (dd, $J = 8.8$, 2.0 Hz, 1H), 6.00 (dd, $J = 8.2$, 2.2 Hz, 1H), 5.38 (d, $J = 12.0$ Hz, 1H), 4.19 (dd, $J = 9.0$, 5.4 Hz, 1H), 3.81 (s, 1H), 3.05–3.25 (m, 4H), 2.38–2.46 (m, 2H), 1.98–2.29 (m, 11H), 1.26–1.74 (m, 15H), 0.99 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 176.2, 176.0, 174.0, 132.2, 130.7, 127.1, 127.0, 126.9, 117.5, 116.6, 116.5, 105.0, 99.8, 84.6, 79.6, 71.5, 69.8, 56.5, 54.4, 37.8, 37.6, 36.1, 32.4, 32.3, 32.0, 30.1, 29.2, 27.5, 26.5, 26.2, 26.1, 25.9, 20.1, 19.7, 18.9, 18.8; LRMS(ESI⁺) m/z calcd for $\text{C}_{37}\text{H}_{46}\text{ClN}_4\text{O}_6$ [M+H]⁺ 692.32; Found 692.28.

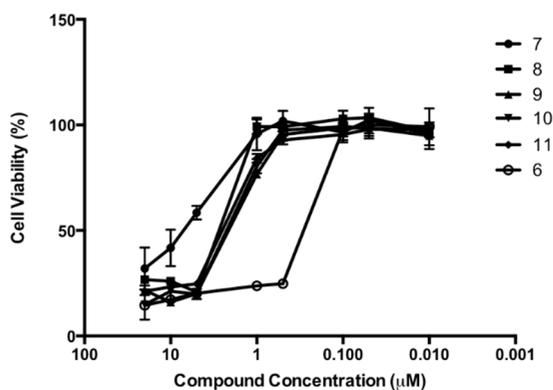
■ (2*S*)-2-(3-(3-(4-(4-chloro-2-methylphenyl)-8-hydroxy-7-methoxy-2-methyl-2*H*-chromen-2-yl)propanamido)propanamido)-4-(3-methyl-3*H*-diazirin-3-yl)-*N*-(prop-2-yn-1-yl)butanamide (**11**)



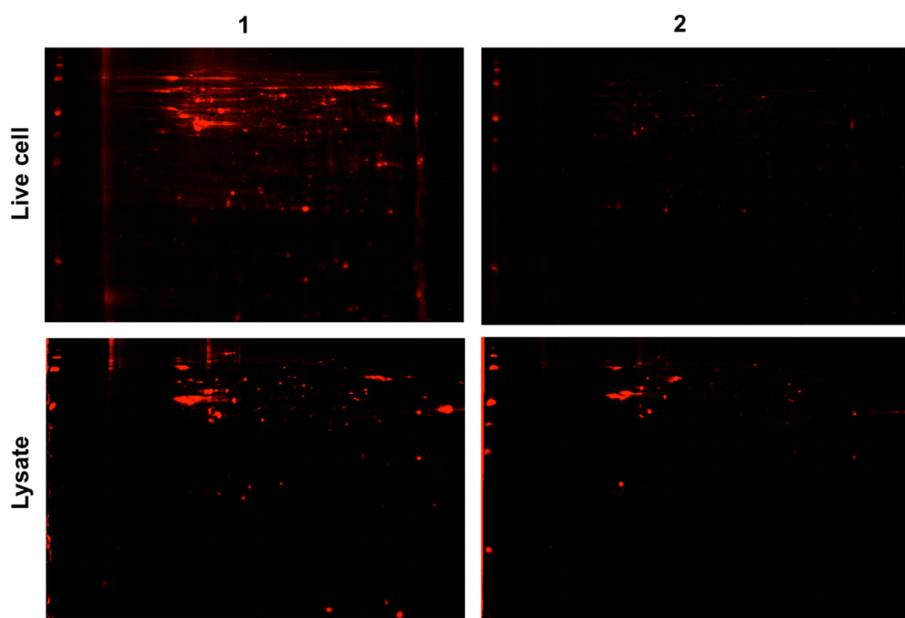
$R_f = 0.62$ (DCM:methanol = 10:1, v/v); $^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.26 (d,

$J = 5.6$ Hz, 1H), 7.22 (d, $J = 8.0$ Hz, 1H), 7.10 (d, $J = 7.6$ Hz, 1H), 6.40 (d, $J = 6.8$ Hz, 1H), 6.00 (d, $J = 7.2$ Hz, 1H), 5.38 (d, $J = 13.2$ Hz, 1H), 4.23 (dd, $J = 8.8, 5.6$ Hz, 1H), 3.85–3.98 (m, 2H), 3.81 (s, 3H), 3.36–3.50 (m, 2H), 2.57 (dd, $J = 4.2, 2.6$ Hz, 1H), 1.98–2.50 (m, 10H), 1.23–1.74 (m, 9H), 0.98 (s, 3H); LRMS(ESI⁺) m/z calcd for C₃₃H₃₈ClN₅O₆ [M+H]⁺ 636.26; Found 636.28.

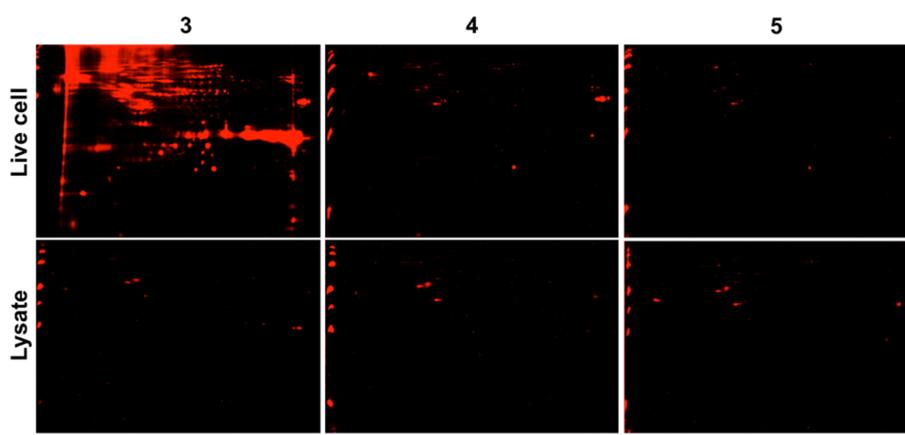
2.3.6. Supporting Information



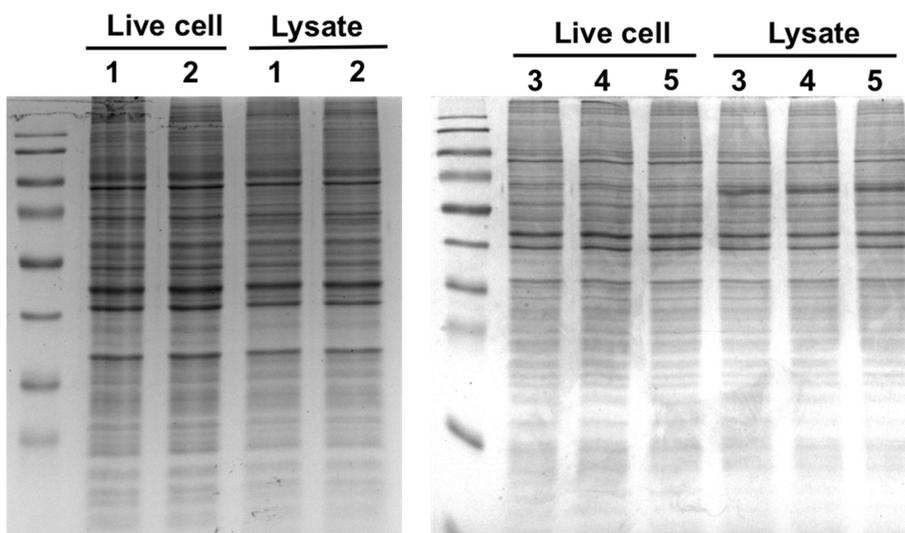
Supplementary Figure S2.3.2. Cell viability against tubulin inhibitor **6** and its photoaffinity probes **7**, **8**, **9**, **10**, and **11**. All the photoaffinity probes have similar range of anticancer activity.



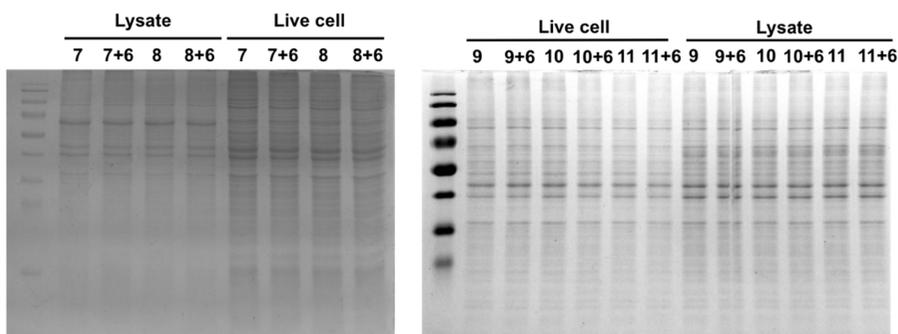
Supplementary Figure S2.3.3. The proteome labeled by PLs **1** and **2** in living cells and cell lysates was subjected to 2D-gel electrophoresis. (**1** and **2**: 10 μ M, HeLa cell).



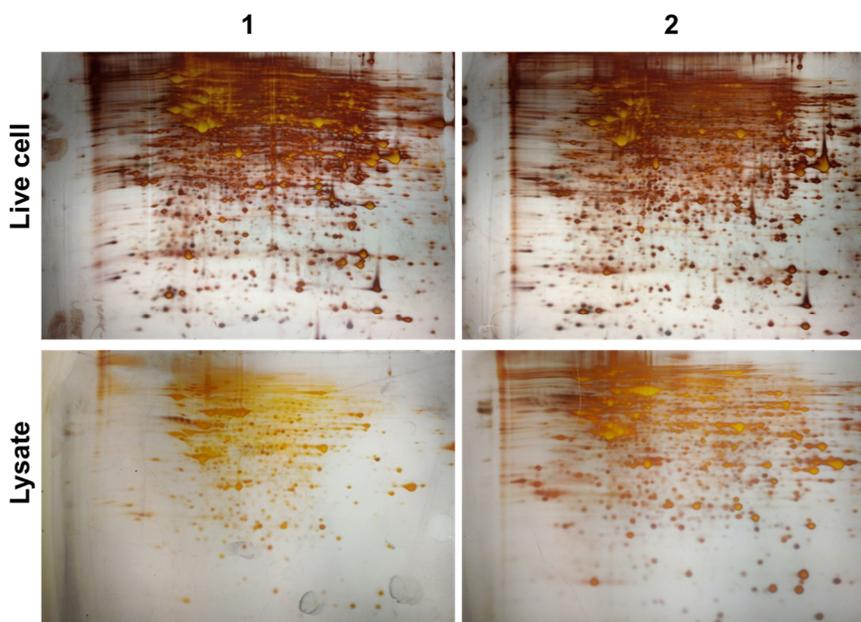
Supplementary Figure S2.3.4. The proteome labeled by PLs **3**, **4**, and **5** in living cells and cell lysates was subjected to 2D-gel electrophoresis. (**3**, **4**, and **5**: 10 μ M, HeLa cell).



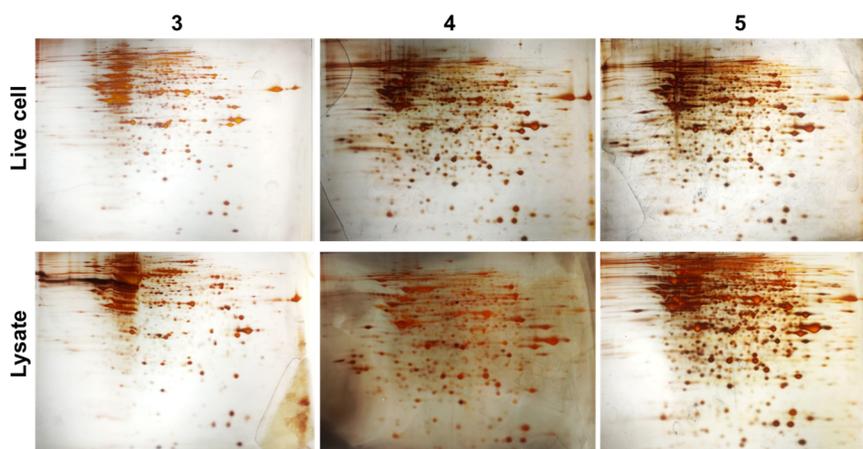
Supplementary Figure S2.3.5. Coomassie staining of 1D gel image in Figure 2.3.1d. Expression patterns of proteomes were not changed by proteome labeling by photoaffinity linkers, but the proteome-labeling pattern was affected by the molecular shape of photoaffinity linker in a structure-dependent manner.



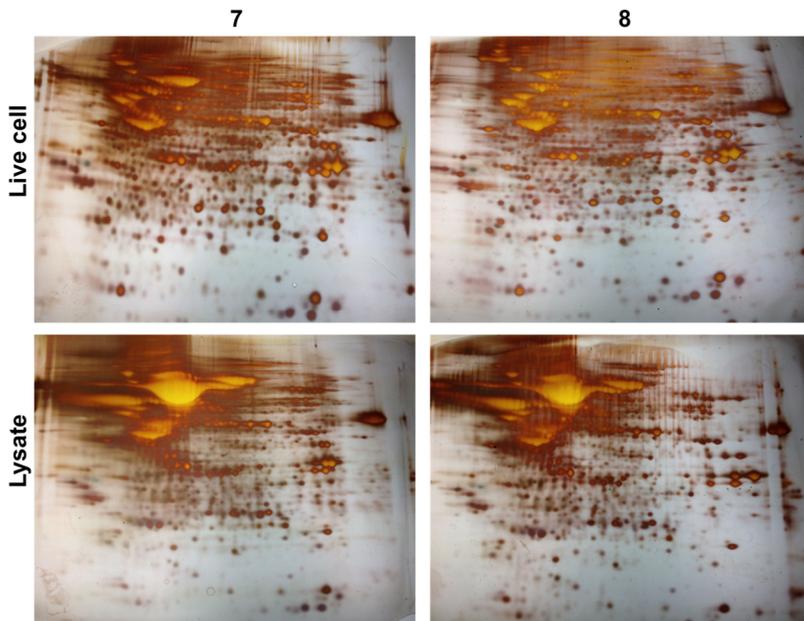
Supplementary Figure S2.3.6. Coomassie staining of 1D gel image in Figure 2.3.2b and 2.3.3b. Expression pattern of proteomes was not changed by proteome labeling upon treatment with probes.



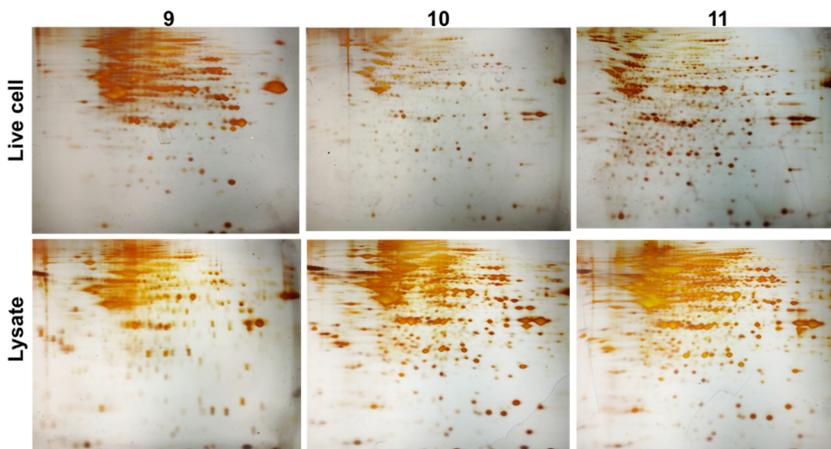
Supplementary Figure S2.3.7. Silver staining of 2D gel image in Figure S2.3.2. Whole proteome expression pattern was not affected by proteome labeling by photoaffinity linkers. Only few proteins out of whole proteome are labeled by each photoaffinity linkers.



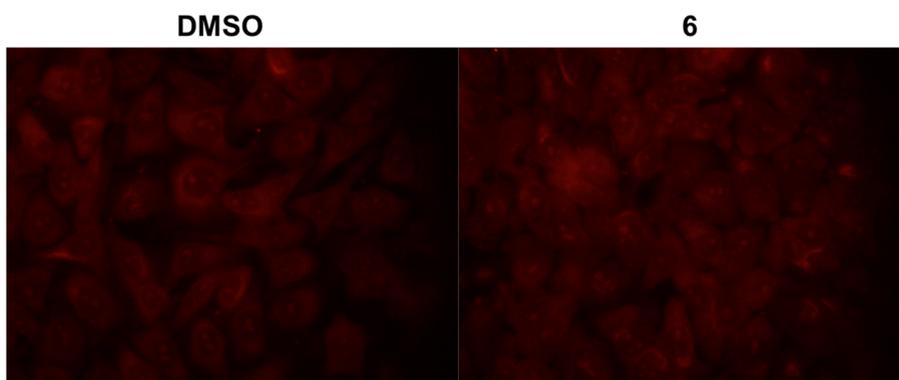
Supplementary Figure S2.3.8. Silver staining of 2D gel image in Figure S2.3.3. Whole proteome expression pattern was not affected by proteome labeling by photoaffinity linkers. Only few proteins out of whole proteome are labeled by each photoaffinity linkers.



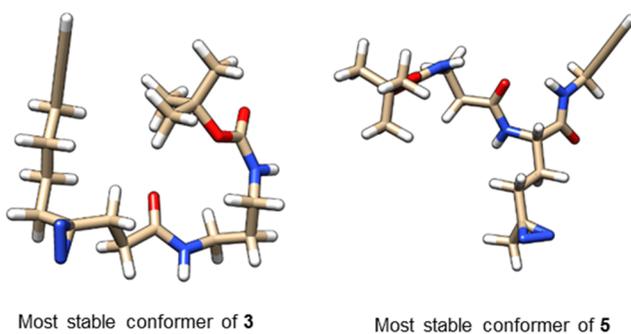
Supplementary Figure S2.3.9. Silver staining of 2D gel image in Figure 2.3.2c. Whole proteome expression pattern was not affected by proteome labeling by photoaffinity linkers. Only few proteins out of whole proteome are labeled by each photoaffinity linkers.



Supplementary Figure S2.3.10. Silver staining of 2D gel image in Figure 2.3.3c. Whole proteome expression pattern was not affected by proteome labeling by photoaffinity linkers. Only few proteins out of whole proteome are labeled by each photoaffinity linkers.



Supplementary Figure S2.3.10. Immunostaining of vimentin structure. Treatment of **6** does not affect the vimentin cellular structures.



Supplementary Figure S2.3.11. Lowest-energy conformation of PL **3** (linear) and **5** (branched)

Cpd.	ClogP	P.E.	< 1	< 2
1	4.0	42.8	26	40
2	3.8	25.4	9	23
3	2.4	-3.6	36	68
4	2.1	-4.4	5	20
5	1.0	-4.3	5	13

Supplementary Figure S2.3.12. cLogP, the minimum energy of the most stable conformer, the number of conformers of **1**, **2**, **3**, **4**, and **5** within a range of 1 or 2 kcal/mol. cLogP value was calculated by ChemBioDraw Ultra 14.0. Energy minimization and structural optimization were

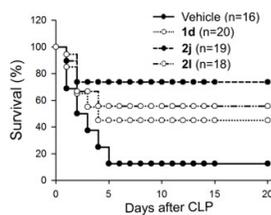
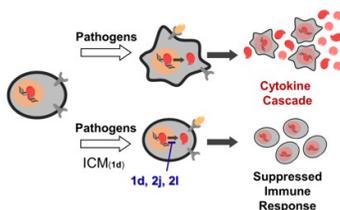
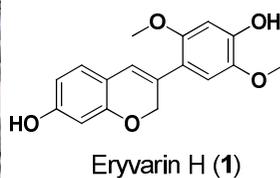
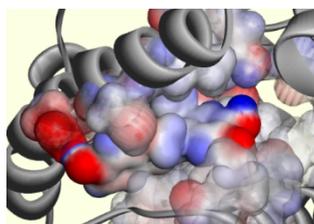
performed by Gaussian W09, potential energies of compounds were calculated by Vconf2.0 from previous optimized structures. Number of conformers within 1 or 2 kcal mol⁻¹ from energy minimum was calculated by Vconf 2.0 based on Tork conformational analysis method.¹⁵

2.3.7. References and Notes

- † Portions of this chapter have been previously reported, see: Park, H.‡; Koo, J. Y.‡; Srikanth, Y. V. V.; Lee, J.; Park, J.; Park, S. B. *Chem. Commun.* **2016**, 52, 5828–5831. (‡: equal contribution)
- (1) (a) Leslie, B. J.; Hergenrother, P. J. *Chem. Soc. Rev.* **2008**, 37, 1347–1360. (b) Su, Y.; Ge, J.; Zhu, B.; Zheng, Y.-G.; Zhu, Q.; Yao, S. Q. *Curr. Opin. Chem. Biol.* **2013**, 17, 768–775. (c) Schenone, M.; Dancik, V.; Wagner, B. K.; Clemons, P. A. *Nat. Chem. Biol.* **2013**, 9, 232–240.
- (2) Sato, S.; Murata, A.; Orihara, T.; Shirakawa, T.; Suenaga, K.; Kigoshi, H.; Uesugi, M. *Chem. Biol.* **2011**, 18, 131–139.
- (3) (a) Park, J.; Koh, M.; Park, S. B. *Mol. BioSyst.* **2013**, 9, 544–550. (b) Sato, S.; Murata, A.; Shirakawa, T.; Uesugi, M. *Chem. Biol.* **2010**, 17, 616–623. (c) Rix, U.; Superti-Furga, G. *Nat. Chem. Biol.* **2009**, 5, 616–624.
- (4) (a) Yang, P. Y.; Liu, K.; Ngai, M. H.; Lear, M. J.; Wenk, M. R.; Yao, S. Q. *J. Am. Chem. Soc.* **2010**, 132, 656–666. (b) Smith, E.; Collins, I. *Future Med. Chem.* **2015**, 7, 159–183.
- (5) Kotzyba-Hibert, F.; Kapfer, I.; Goeldner, M. *Angew. Chem. Int. Ed.* **1995**, 34, 1296–1312.
- (6) (a) Li, Z.; Hao, P.; Li, L.; Tan, C. Y. J.; Cheng, X.; Chen, G. Y. J.; Sze,

- S. K.; Shen, H.-M.; Yao, S. Q. *Angew. Chem. Int. Ed.* **2013**, *52*, 8551–8556. (b) Li, Z.; Wang, D.; Li, L.; Pan, S.; Na, Z.; Tan, C. Y. J.; Yao, S. Q. *J. Am. Chem. Soc.* **2014**, *136*, 9990–9998.
- (7) (a) Cisar, J. S.; Cravatt, B. F. *J. Am. Chem. Soc.* **2012**, *134*, 10385–10388. (b) Kambe, T.; Correia, B. E.; Niphakis, M. J.; Cravatt, B. F. *J. Am. Chem. Soc.* **2014**, *136*, 10777–10782.
- (8) (a) Li, G.; Liu, Y.; Liu, Y.; Chen, L.; Wu, S.; Liu, Y.; Li, X. *Angew. Chem. Int. Ed.* **2013**, *52*, 9544–9549. (b) Li, G.; Liu, Y.; Yu, X.; Li, X. *Bioconjugate Chem.* **2014**, *25*, 1172–1180.
- (9) Niphakis, M. J.; Lum, K. M.; Coggnetta III, A. B.; Correia, B. E.; Ichu, T.-A.; Olucha, J.; Brown, S. J.; Kundu, S.; Piscitelli, F.; Rosen, H.; Cravatt, B. F. *Cell* **2015**, *161*, 1668–1680.
- (10) Zhou, H.; Rivas, G.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375–397.
- (11) Park, J.; Oh, S.; Park, S. B. *Angew. Chem. Int. Ed.* **2012**, *51*, 5447–5451.
- (12) Kim, Y. K.; Lee, J.-S.; Bi, X.; Ha, H.-H.; Ng, S. H.; Ahn, Y.-H.; Lee, J.-J.; Wagner, B. K.; Clemons, P. A.; Chang, Y.-T. *Angew. Chem. Int. Ed.* **2011**, *50*, 2761–2763.
- (13) Park, J.; Koh, M.; Koo, J. Y.; Lee, S.; Park, S. B. *ACS Chem. Biol.* **2016**, *1*, 44–52.
- (14) Shi, H.; Zhang, C. J.; Chen, G. Y. J.; Yao, S. Q. *J. Am. Chem. Soc.* **2012**, *134*, 3001–3014.
- (15) Chang, C.-E.; Gilson, M. K. *J. Comput. Chem.* **2003**, *24*, 1987–1998.

Part 3. Structural Optimization of Hit Compounds based on Identified Biological Targets



3.1. Total Synthesis of Eryvarin H and its Derivatives as Estrogen-related Receptor γ (ERR γ) Inverse Agonists

3.1.1. Abstract

Total synthesis of eryvarin H and the biological investigation of its analogues as a potential inverse agonist of ERR γ are depicted here. Among 13 analogues prepared by modular synthetic route, eryvarin H and compound **12** showed meaningful ERR γ inverse agonistic activities along with moderate selectivity over ER α and other nuclear receptors in cell-based reporter gene assay.

3.1.2. Introduction

Nuclear receptors (NRs) have been recognized as one of the most important targets for the development of novel therapeutic agents via regulating various disease-relevant cellular functions at the transcription level.¹ NRs are described as transcriptional factors mainly regulated by endogenous ligands in biological systems. Among them, a series of NRs are categorized as orphan nuclear receptors when they don't have any identified endogenous ligands, and some of orphan nuclear receptors still possess constitutive transcriptional activities without their endogenous ligands.² Therefore, the identification of bioactive small molecules that can specifically control their activity can provide an

important clue for the development of novel therapeutic agents as well as a research tool for deciphering complex events in biological systems.³ Unlike genetic approach, bioactive small molecules can perturb particular functions of wild-type gene products in a temporal and reversible manner.⁴ Therefore, we aimed to identify new synthetic ligands for the transcriptional regulation of nuclear receptors, especially orphan nuclear receptors.

We have been interested in estrogen-related receptors (ERRs) that are orphan NRs and closely related to estrogen receptors (ERs) in their sequence homology. But their difference in the ligand binding domain (LBD) differentiate their characteristics from those of ERs.⁵ ERRs have various functional roles associated with diseases including cancer and metabolic diseases as ER does.⁶ Among members of ERR family, we particularly paid an attention to ERR γ due to its crucial roles in various biological events, such as hepatic insulin signaling, gluconeogenesis, regulating oxidative metabolism, suppressing tumor growth of prostate cancer cells, and modulating cell proliferation and estrogen signaling in breast cancer.⁶⁻⁸ ERR γ is a third subtype receptor and highly expressed in various human adult tissues including brain, kidney, skeletal muscle, heart, and placenta.⁸ Based on these observations, ERR γ became an emerging target for cancers and metabolic diseases, even though there are a limited number of ERR γ agonists and inverse agonists that control its downstream activity.¹⁰⁻¹² In addition, most of reported ligands suffered non-specific interactions with various gene products and caused functional cross-talks with other nuclear receptors, especially with ER α , in the estrogenic signaling pathway.^{10,11} Therefore, it is essential to discover novel bioactive

small molecules that selectively regulate ERR γ to elucidate its functional roles in our body. As a continuation of our efforts^{7,10} in this field, we pursued the identification of novel ERR γ inverse agonists to deliver new potential drug-like molecules to the biomedical community.

Under this mission statement, we searched the new chemical entity from the collection of natural products using structure-based *in silico* analysis. With x-ray co-crystal structure of ERR γ with GSK5182, a known ERR γ inverse agonist, we performed a docking simulation study of natural product collection at the LBD of ERR γ , which allows the identification of potential ligands for ERR γ . After computer-based screening of 4000 structurally diverse natural products from Korea Bioactive Natural Material Bank, we selected natural product eryvarin H (**1**) as a potential inverse agonist of ERR γ . Eryvarin H was first isolated from the roots of *Erythrina* species, *Erythrina variegata*¹³ or *Erythrina abyssinica* in 2003.¹⁴ In the past, eryvarin H has been known as an antimicrobial agent¹⁴ or as a cellular radical scavenger.¹⁵ However, eryvarin H (**1**) used in previous reports was extracted from natural resources with a limited quantity. Therefore, herein, we accomplished the first total synthesis of eryvarin H (**1**) and its derivatives using a modular synthetic route. The resulting eryvarin H and its derivatives were evaluated for their function as selective ligands for ERR γ .

3.1.3. Results and Discussion

For the discovery of novel ERR γ inverse agonist, we first explored the *in silico* docking study of structurally diverse natural products with ERR γ co-

crystalized with GSK5182 (PDB ID: 2GPU).¹² GSK5182 is a confirmed ERR γ inverse agonist and binds at its ligand binding site. From the collection of natural products, we searched for potential ligands that can bind at the LBD of ERR γ in a similar way of GSK5182. Based on the binding scores in computational docking analysis, we selected a list of natural products. Among them, we selected eryvarin H (**1**) as a potential inverse agonist of ERR γ . As shown in Figure 3.1.1A–B, eryvarin H (**1**) nicely occupies the empty space at the active site of ERR γ with dipole-dipole interaction as well as hydrophobic interactions, which was clearly visualized in the electrostatic potential clouds of eryvarin H at the LBD of ERR γ . The whole amino acid map within the range of 2 Å also reveals the specific interactions, such as hydrogen bonding of Asp273, Tyr326 and Asn346 with hydroxyl groups on eryvarin H (**1**) and dipole-dipole interaction of Cys269 with ether linkage on eryvarin H (**1**) (Figure 3.1.1C). However, even though there is a series of potential specific interactions of eryvarin H with ERR γ , it is essential to have a structural modularity to ensure the specificity due to the high degree of structural similarity of ERR γ with ER α . In addition, the extraction from natural resources only provides the limited quantity of eryvarin H without an access to its analogues. Therefore, we pursued the total synthesis of eryvarin H (**1**) and the subsequent structural modification for structure-activity relationship study.

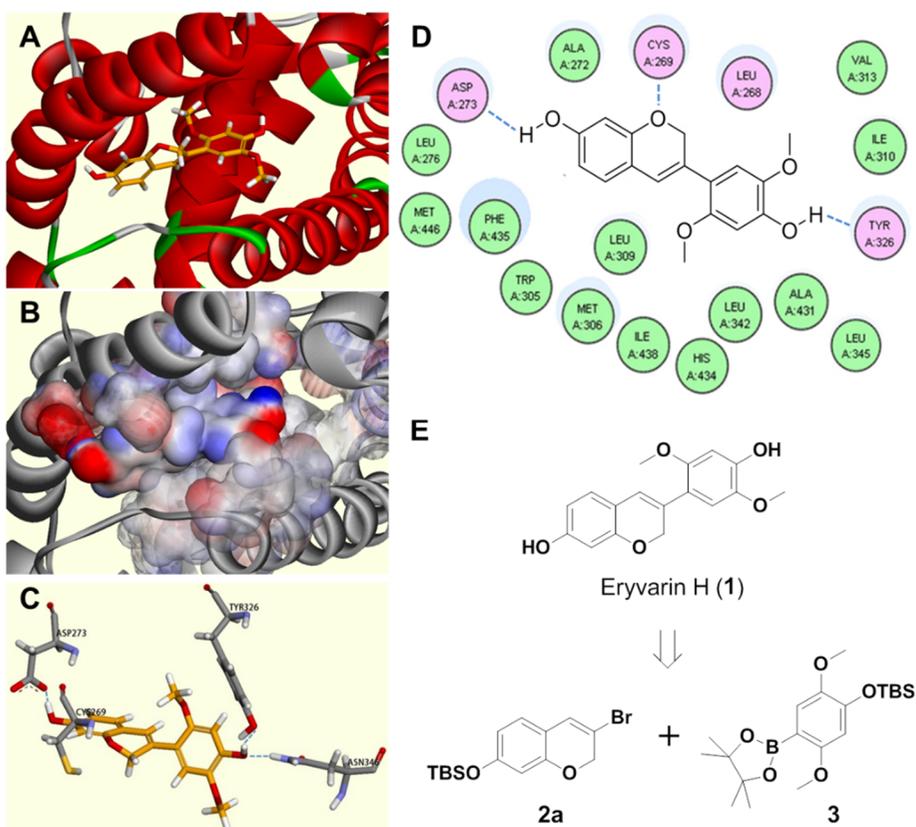
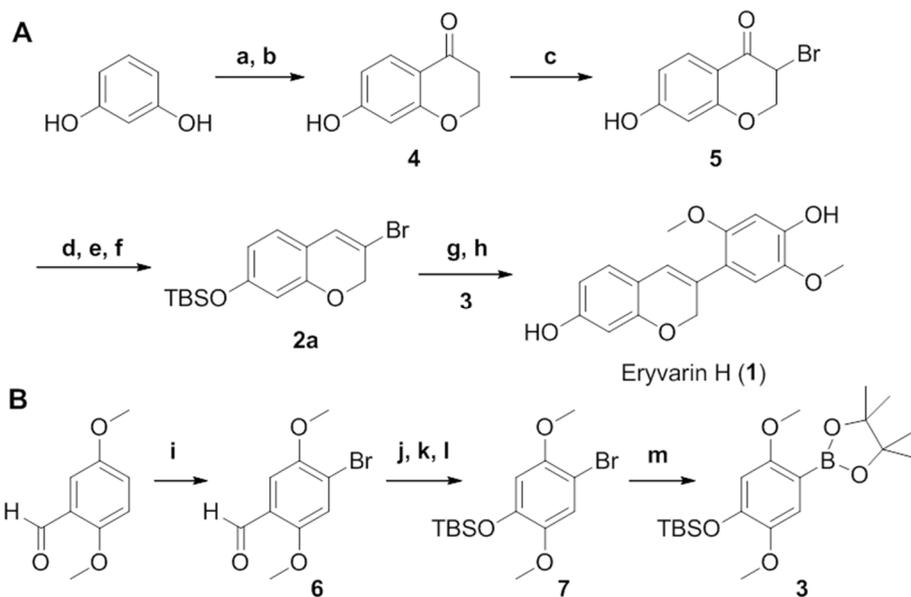


Figure 3.1.1. Docking simulation of eryvarin H at the ligand binding pocket of ERR γ . (A) Potential binding mode of eryvarin H; (B) Possible hydrophobic and dipole-dipole interactions of eryvarin H with residues at the ligand binding site of ERR γ shown in space filling model. The red color represents partial positive charge, and the blue color represents partial negative charge; (C) Potential hydrogen bonding interactions of eryvarin H with polar amino acids (Asp273, Tyr326, and Asn346); (D) Schematic diagram of amino acids within 2 Å range from eryvarin H; (E) Chemical structure of eryvarin H and its retrosynthetic analysis.

Eryvarin H (**1**) is composed of two aromatic rings, 2H-chromen-7-ol and electron-rich aryl ring, that can be connected via C-C bond formation using Pd-mediated cross-coupling (Figure 3.1.1E). Based on this retrosynthetic analysis of eryvarin H (**1**), we prepared the two cross-coupling partners, vinyl bromide

2a and arylboronic ester **3**. For the preparation of vinyl bromide (**2a**) part, resorcinol was first transformed to 7-hydroxychroman-4-one (**4**) by treatment of 3-chloropropionic acid via acid-catalyzed electrophilic aromatic acylation, followed by *O*-alkylation.¹⁶ The mono-bromination at the α position to carbonyl in compound **4** led to the formation of compound **5**. The *tert*-butyldimethylsilyl (TBS) protection of phenol in compound **5**, carbonyl reduction using NaBH₄, and the subsequent acid-catalyzed dehydration allowed the formation of desired vinyl bromide (**2a**) in good yields (Scheme 3.1.1A). The synthesis of its cross-coupling counterpart, arylboronic ester (**3**), was started by the treatment of 2,5-dimethoxybenzaldehyde with Br₂ under mild acidic condition to provide its brominated compound **5**.¹⁷ This aryl bromination was quite regioselective at the C-4 position due to the presence of aldehyde moiety as a biasing element, which was interconverted to phenol moiety through Baeyer-Villiger oxidation of **6** and the direct hydrolysis of resulting ester in basic methanol solution. The resulting phenol was then protected by TBSCl to prepare **7** in excellent 3-step yields. Finally, various conditions were tested for the introduction of C-C cross-coupling moiety, but Pd-catalyzed boronic ester formation with pinacolborane, DPEphos, triethylamine (TEA) in 1,4-dioxane only yielded the desired product **3** in a reasonable yield (Scheme 3.1.1B). After the preparation of both key partners, we successfully finished up the first total synthesis of eryvarin H (**1**) via the formation of C-C covalent bond between vinyl bromide **2a** and arylboronic ester **3** using Suzuki-Miyaura coupling reaction, followed by TBS deprotection with HF/pyridine with 8 linear steps in 18% overall yield (see Supporting Information).



Scheme 3.1.1. Synthetic route for total synthesis of eryvarin H via C-C bond formation between two coupling partner (2a and 3). *Reagents and conditions:* (A) Preparation of eryvarin H (**1**) through **2a**. (a) 3-Chloropropionic acid, TfOH, 80 °C, 2 h; (b) 2N NaOH, 0 °C to r.t., 3 h, 62% (2-step yield); (c) CuBr₂, EtOAc/CHCl₃/MeOH, 70 °C, 4 h, 67%; (d) TBSCl, Imidazole, DCM, r.t., 1 h; (e) NaBH₄, EtOH, r.t., 1 h; (f) TsOH-H₂O, toluene, 80 °C, 120 W, 20 min, 61% (3-step yield); (g) **3**, Pd(PPh₃)₄, Na₂CO₃, toluene/EtOH/H₂O, 80 °C, 2 h, (h) HF/pyridine, THF, r.t., 3 h, 73% (2-step yield). (B) Preparation of **3**. (i) Br₂, acetic acid, r.t., 12 h, 62 %; (j) *m*-CPBA, DCM, r.t., 3 h; (k) NaOH, methanol, r.t., 3 h; (l) TBSCl, imidazole, DCM, r.t., 3 h, 95% (3-step yield); (m) Pinacolborane, Pd(OAc)₂, DPEphos, TEA, 1,4-dioxane, 100 °C, 12 h, 63%.

After the completion of total synthesis of eryvarin H, we envisioned the preparation of its systematic analogues for the molecular level of understanding of its binding event with ERR γ . On the basis of our *in silico* docking simulation

of eryvarin H at the LBD of ERR γ , we hypothesized that the specific interaction of 2',5'-dimethoxy-4'-hydroxyphenyl on eryvarin H at the deep binding pocket of ERR γ is crucial for its functional modulation. We also revealed some flexible and empty regions at the ligand binding site in the docking structure of ERR γ with eryvarin H, which might provide an opportunity to induce some additional binding interactions for enhanced potency and selectivity toward ERR γ over ER α . Therefore, we prepared 12 derivatives (**8–19**) of eryvarin H using the modular synthetic procedure (see Supporting Information). As shown in Figure 3.1.2, we focused on the diversification of aryl substituents to decipher the origin of specific hydrogen bondings at the ligand binding pocket. Dimethyl-substituted vinyl bromide **2b**, TIPS-protected 3-bromo-2,2-dimethyl-2*H*-chromen-7-ol, was synthesized using a reported procedure.¹⁸

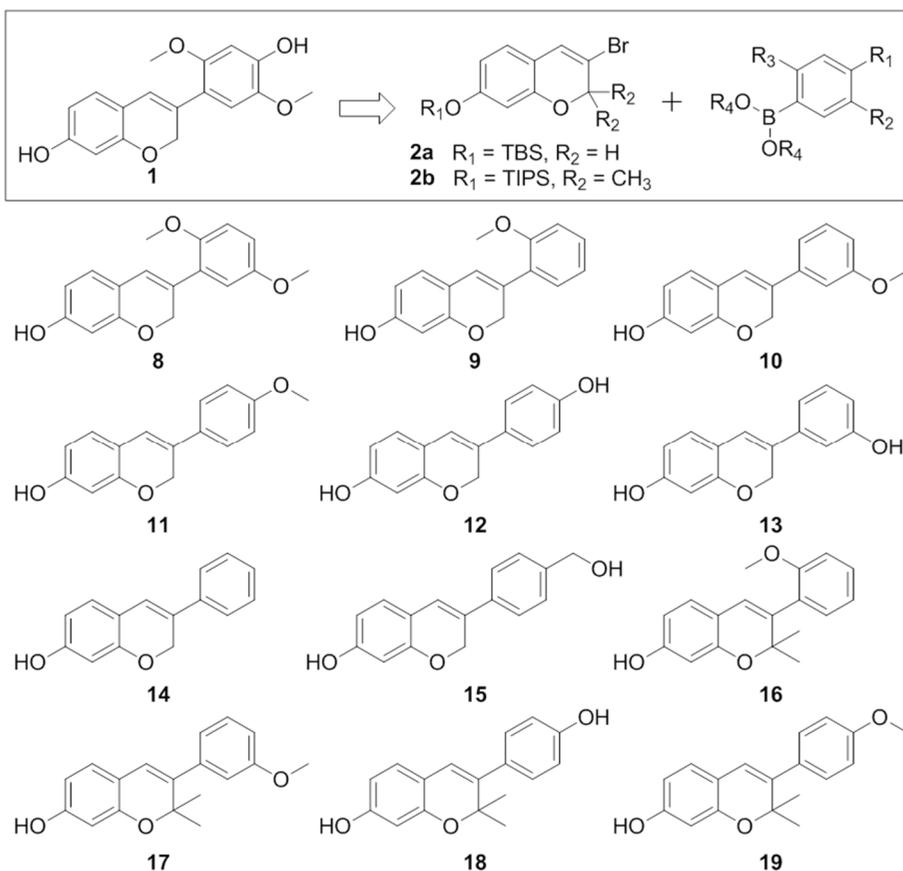


Figure 3.1.2. Derivatives of eryvarin H containing isoflav-3-ene structural motif (8–19). These derivatives were systematically prepared by Pd-mediated Suzuki cross-coupling between aryl bromide structures (**2a** or **2b**) and various boronic esters/boronic acids.

After the preparation of eryvarin H (**1**) and its analogues (**8–19**), their biological activities on the transcriptional regulation toward ERR γ were evaluated using cell-based reporter gene assay to confirm their roles as inverse agonists. In this assay, we measure the chemoluminescence induced by the expression changes of luciferase upon treatment of small molecules after the transient transfection of DNA plasmid (Gal4-fused ERR γ -LBD construct) to

HEK-293T human kidney cell line by calcium phosphate transfection protocol.¹⁹ The resulting luminescence signal was normalized with β -gal expression to minimize the false positives caused by cellular cytotoxicity.²⁰ As shown in Figure 3.1.2, we confirmed that eryvarin H (**1**) is an ERR γ inverse agonist and showed a drastic reduction in the chemoluminescence signal of downstream luciferase. Interestingly, all 13 analogues including eryvarin H showed some levels of inverse agonistic activities toward ERR γ , though the irrelative luminescence signals were varied on the basis of chemical structure (see ESI†). Among them, compounds **1**, **11**, and **12** showed more potent inverse agonism. Speaking of their structure-activity relationship, either hydroxy or methoxy group at the C-4' position is essential for their activity and the steric or hydrophobic elements are not necessary at the C-2 position in the case of analogues (**16–19**) containing 2*H*-chromenol with dimethyl substituents instead of two hydrogens. The additional bulky dimethyl substituents at the C-2 position might block the specific interactions with ligand binding pocket of ERR γ , which significantly reduced the inverse agonistic activity.

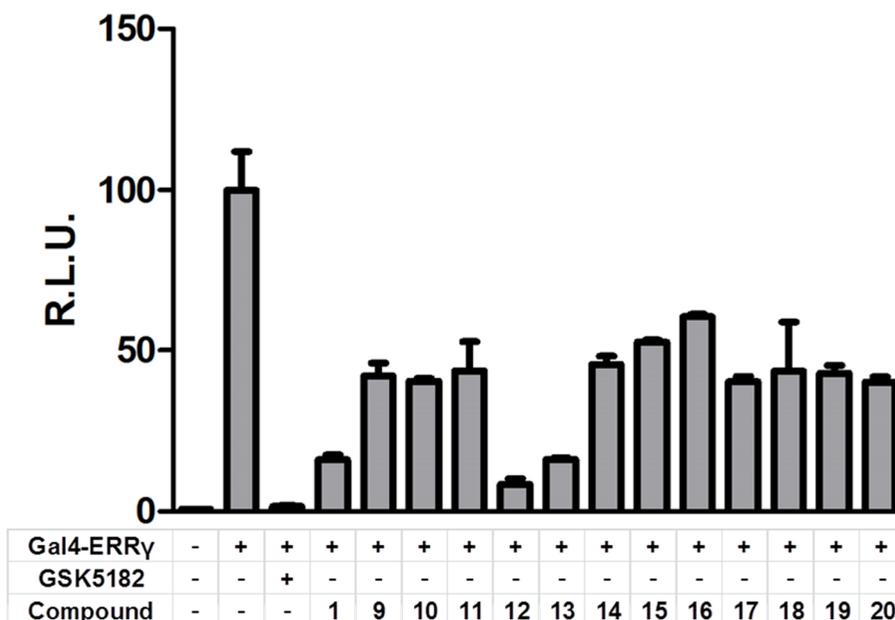


Figure 3.1.3. Biological evaluation of eryvarin H (1) and its derivatives (8–19) as inverse agonists using cell-based reporter gene assay in HEK-293T cells after transfection of Gal4-ERR γ , pFR-Luc, and β -gal plasmids. Compounds 1, 11 and 12 showed good inverse agonistic activities among all tested compounds. All compounds including GSK5182 were treated at the final concentration of 10 μ M. R.L.U. is relative luminescent unit.

Furthermore, transcriptional activities of these 13 analogues towards other kinds of NRs (ER α , mCAR, HNF4, SF-1)²¹ were measured by reporter gene assay systems to investigate the selective ERR γ inverse agonistic activity over other NRs (Figure 3.1.4). Owing to the high degree of structural similarity and DNA sequence homology of ERR γ with ER α , the ligand-induced transcriptional activity of ERR γ is compared with that of ER α . As shown in Figure 3.1.3 and 3.1.4, eryvarin H and compound 12 showed the selective inverse agonistic activity toward ERR γ over ER α . However, the efficacy and

specificity of eryvarin H we observed in this study are not comparable to those of GSK5182, that we have been used for the biochemical studies of ERR γ .

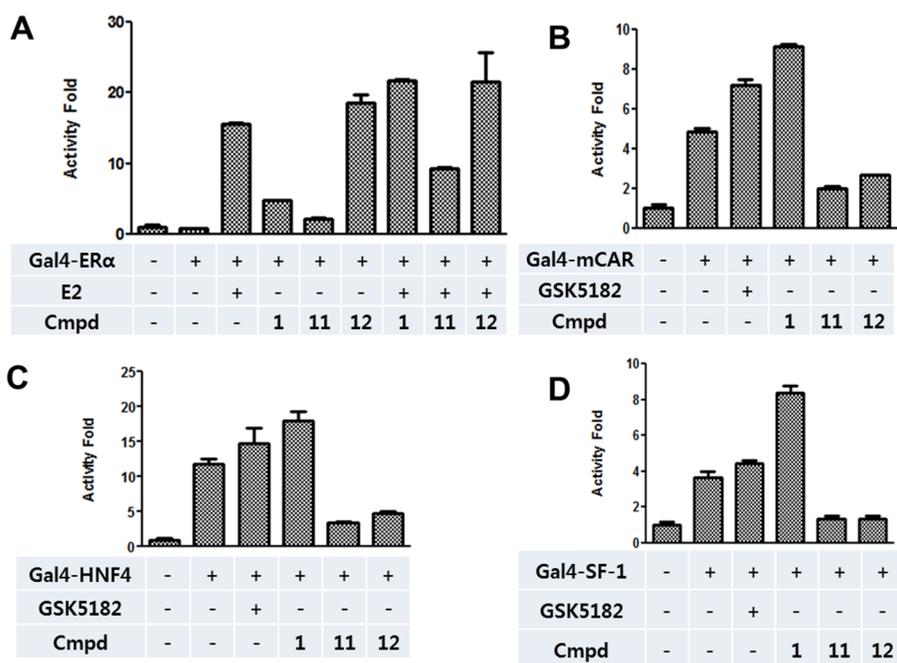


Figure 3.1.4. Transcriptional activities of eryvarin H, 11, and 12 towards other NRs. (a) ER α , (b) mCAR, (c) HNF4 and (d) SF-1. Normalized luciferase activities of each NR were tested in HEK-293T cells upon treatment of three compounds to test their selective inverse agonism over that in ERR γ . E2 refers to estradiol.

3.1.4. Conclusion

In conclusion, natural products have been serving as essential perturbagens in various biological processes due to their evolutionary heritage and structural diversity. As a continuation of our effort on the functional modification of

ERR γ , we aimed to develop a novel ERR γ inverse agonist via *in silico* docking analysis with a collection of natural products. The computer-aided ligand discovery allowed the identification of eryvarin H (**1**) as a potential inverse agonist of ERR γ . Along with a limited supply of eryvarin H via extraction from natural resources, the inability of structure-activity relationship of this natural product against ERR γ lead us to pursue the first total synthesis of eryvarin H and its structural analogues. In addition, we wanted to decipher the molecular basis of its interaction with residues at the ligand binding pocket of ERR γ , which has high structural similarity and sequence homology with ER α . Therefore, we prepared two key intermediates, vinyl bromides (**2a**) and arylboronic ester (**3**), and successfully completed the total synthesis of eryvarin H through Pd-mediated Suzuki-Miyaura cross-coupling of two intermediates in reasonable overall yields. With this modular synthetic route, we also prepared 12 derivatives (**8–19**) of eryvarin H simply by changing the substituents on arylboronic esters or introducing a dimethyl substituent at the C-2 position of TBS-protected 3-bromo-2*H*-chromen-7-ol. The resulting 13 analogues including eryvarin H were subjected to cell-based reporter gene assay using DNA plasmid containing Gal4-fused ERR γ LBD to measure their biological activity as ERR γ inverse agonists. Among these derivatives, eryvarin H and compound **12** showed meaningful ERR γ inverse agonistic activities along with moderate selectivity over ER α and other NRs. However, the level of efficacy and specificity of eryvarin H toward ERR γ over ER α is not comparable to that of GSK5182, that we have been used for the biochemical studies of ERR γ . Therefore, we unfortunately discontinued the follow-up *in vivo* biological study

of eryvarin H and its analogues due to their limited potential as a selective ERR γ inverse agonist. Even though this study failed to produce a selective and potent inverse agonist of ERR γ , this study clearly exemplified the rational drug discovery procedure using *in silico* docking simulation with a collection of natural product, followed by total synthesis of a natural product and its analogues as a potential perturbagen of protein-of-interests. We envision that this rational approach can be adapted to identify new small molecule that modulates the function of new therapeutic targets and key signaling pathways.

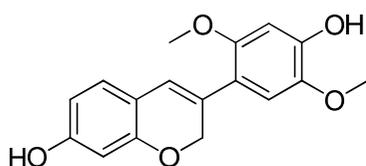
3.1.5. Experimental Section

Cell-based Reporter Gene Assay

The HEK-293T cells were seeded into 24-well plates at a density of 2.0-8.0 x 10⁴ cells/well 24 h prior to the transfection. The cells were transiently transfected with pFR(5xGal4 binding site)-Luc, pCMX-Gal4-ERR γ , ER α , mCAR, HNF α , SF-1 and pCMV- β -gal for the ERR γ reporter gene assay. Transient transfection was performed using SuperFect (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The cells were treated with 100nM E2(estradiol), 10 μ M GSK5182, 10 μ M prepared compounds (**1**, **8**–**19**) for the final 24 h. The cells then were harvested 48h after transfection, and luciferase activity was measured. Luciferase activity was normalized to β -galactosidase activity. The data is representative of at least three to five independent experiments.

Synthetic Procedures of Compound 1–19.

Compound **1**, **Eryvarin H**, 3-(4-hydroxy-2,5-dimethoxyphenyl)-2H-chromen-7-ol

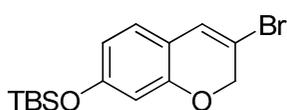


Compound **2a** (0.100 mmol), one of boronic acids/boronic esters (2 equiv.), Pd(PPh₃)₄ (5 mol %), Na₂CO₃ (3 equiv.) were suspended in a solvent mixture of Toluene/EtOH/H₂O (2:1:1, total volume of 4 mL). The reaction mixture was stirred at 80 °C for 4 h and the completion of the reaction was monitored by TLC. When the starting material **2a** was all consumed, the resulting mixture was diluted with ethyl acetate and washed with brine. The combined organic layer was dried over anhydrous MgSO₄, then filtered, and concentrated under reduced pressure. After a short silica-gel filtration, the resulting product was used directly for the next reaction without further purification. The solution of previous prepared compound in HF/pyridine/THF (1:1:18, volumetric ratio, total volume of 3 mL) was stirred for 5 h at room temperature in a plastic vessel. After the reaction completion, Additional fluoride source was quenched with excess TMSOEt (6 mL). The mixture was evaporated under reduced pressure, the resultant was purified directly with silica gel flash column chromatography (EtOAc:hexane = 1:4 to EtOAc:hexane = 1:3) to provide the desired product, compound **1**.

Yield: 73% (2-step yield), *R_f* = 0.04 (1:6 = EtOAc:hexane, v/v); ¹H NMR (500 MHz, CDCl₃) δ 6.93 (d, 1H, *J* = 8.0 Hz), 6.83 (s, 1H), 6.56 (s, 1H), 6.49 (s, 1H), 6.39–6.41 (m, 2H), 4.98 (d, 2H, *J* = 1.0 Hz), 3.88 (s, 3H), 3.75 (s, 3H);

^{13}C NMR (75 MHz, $\text{DMSO-}d_6$) δ 158.2, 154.1, 151.6, 147.5, 141.6, 127.8, 127.5, 120.4, 117.1, 115.3, 113.0, 108.7, 102.4, 101.0, 67.7, 56.5, 55.9; HRMS (FAB $^+$) m/z calcd for $\text{C}_{17}\text{H}_{16}\text{O}_5$ $[\text{M}]^+$ 300.0998, found 300.0996.

Compound **2a**, (3-bromo-2H-chromen-7-yloxy)(tert-butyl)dimethylsilane



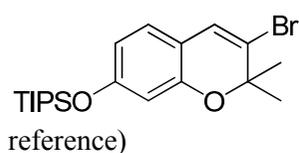
The mixture of **5** (3 mmol, 1 equiv.) and imidazole (1.5 equiv.) was stirred in anhydrous DCM (20 mL) for 15 min at room temperature. To the solution was added TBSCl (1.1 equiv.) and the reaction mixture was stirred for 1 h at room temperature. Then the solution was extracted with excess ethyl acetate and washed with saturated NH_4Cl aqueous solution and brine. The organic layer was dried over anhydrous MgSO_4 and filtered. Then, the filtrate was condensed under reduced pressure. The resulting product was used directly for the next reaction without further purification. To a solution of resulting compound (1 equiv.) in EtOH (20 mL), sodium borohydride (NaBH_4 , 1.2 equiv.) was added at room temperature and the reaction mixture was stirred for 1 h at room temperature. After the completion of reaction monitored by TLC, the resulting mixture was diluted with deionized water and quenched by aqueous NH_4Cl solution, and extracted three times with ethyl acetate. The combined organic layer was washed with brine once, then dried over anhydrous MgSO_4 , filtered, and concentrated under reduced pressure. The resulting compound is α -bromoalcohol structure 3-bromo-7-(tert-butyldimethylsilyloxy)chroman-4-ol as a diastereomeric mixture. To a solution of previous α -bromoalcohol structure (1 equiv.) in anhydrous toluene, p-toluenesulfonic acid monohydrate (p-TSA, 0.1 equiv.)

was added and the reaction mixture was heated in capped microwave vessel under microwave irradiation (80 °C, 120 W) for 20 min. After the reaction completion monitored by TLC, the resulting mixture was diluted with ethyl acetate and washed with brine. The combined organic layer was dried over anhydrous MgSO₄, filtered, and concentrated under reduced pressure. The desired product **2a** was obtained by the purification using silica-gel flash column chromatography (only hexane to EtOAc:hexane = 1:50).

Yield: 61% (3-step yield), *R_f* = 0.43 (Only hexane); ¹H NMR (500 MHz, CDCl₃) δ 6.78 (d, 1H, *J* = 8.0 Hz), 6.68 (s, 1H), 6.37 (dd, 1H, *J* = 2.3, 8.3 Hz), 6.31 (dd, 1H, *J* = 0.5, 2.5 Hz), 4.83 (d, 2H, *J* = 1.5 Hz), 0.96 (s, 9H), 0.19 (s, 6H); ¹³C NMR (125 MHz, acetone-*d*₆) δ 158.0, 154.3, 128.1, 126.8, 117.4, 114.6, 112.7, 108.7, 70.8, 26.2, 19.0, -4.1; HRMS (FAB⁺) *m/z* calcd for C₁₅H₂₁BrO₂Si [M]⁺ 340.0494, found 340.0489.

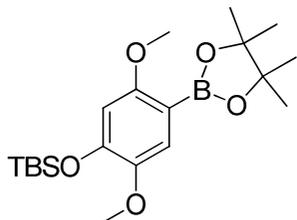
Also, **2a** was obtained from α-bromoalcohol structure by thermal dehydration (80 °C for 2 h) with the 3-step yield of 24%. Stability of **2a** was not quite good, thus it should be kept in -20 °C and more stable as a solution in ethyl acetate or hexane.

Compound **2b**, (3-bromo-2,2-dimethyl-2H-chromen-7-yloxy) triisopropylsilane



Compound **2b** was previously reported in *Chem. Commun.*, 2006, 2962–2964. (Compound **17a** in the

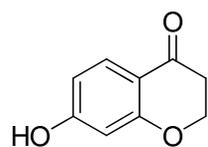
Compound **3**, *tert*-butyl(2,5-dimethoxy-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)dimethylsilane



To the solution of starting material **7** (0.300 mmol, 1 equiv.), Pd(OAc)₂ (5 mol %), DPEphos (10 mol %) in 1,4-dioxane (2 mL) was added triethylamine (4 equiv.) and pinacolborane (3 equiv.). The reaction mixture was stirred at 80 °C for 12 h. After the completion of the reaction monitored by TLC, the resulting mixture was diluted with aqueous NH₄Cl saturated solution, and extracted three times with ethyl acetate. The combined organic layer was washed with brine once, then dried over anhydrous MgSO₄, filtered, and concentrated under reduced pressure. The desired product **3** was obtained by the purification using silica-gel flash column chromatography (EtOAc:hexane = 1:20 to EtOAc:hexane = 1:4).

Yield: 63%, *R*_f = 0.47 (1:6 = EtOAc:hexane, v/v); ¹H NMR (500 MHz, CDCl₃) δ 7.17 (s, 1H), 6.43 (s, 1H), 3.79 (s, 3H), 3.75 (s, 3H), 1.34 (s, 12H), 0.99 (s, 9H), 0.16 (s, 6H); ¹³C NMR (125 MHz, CDCl₃) δ 164.2, 159.9, 149.1, 144.9, 120.5, 105.9, 83.4, 56.8, 56.4, 25.9, 25.0, 18.6, -4.5; HRMS (FAB⁺) *m/z* calcd for C₂₀H₃₅O₅SiB [M]⁺ 394.2351, found 394.2353.

Compound **4**, 7-hydroxychroman-4-one

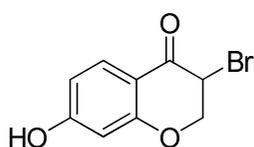


To a mixture of resorcinol (10.0 mmol, 1 equiv.) and 3-chloropropionic acid (1.01 equiv.) was slowly added trifluoromethanesulfonic acid (5 mL). The solution was stirred at 80 °C for 1.5 h, and poured into DCM (100 mL). The solution was

poured into deionized water and the aqueous layer was extracted with DCM (100 mL) twice. The combined organic layer was dried over anhydrous MgSO_4 , then filtered, and concentrated under reduced pressure. The resulting product was used directly for the next reaction without further purification, with addition of 2N NaOH (400 mL) aqueous solution was stirred at 0 °C for 4 h. After the reaction was completed, the pH was adjusted to 2 with concentrated HCl by checked with pH paper. The mixture was extracted trice with EtOAc, washed with brine, dried over MgSO_4 , filtered, and concentrated under reduced pressure. The desired product **4** was obtained by recrystallization with EtOAc/hexane (1:5).

Yield: 62% (2-step yield), $R_f = 0.41$ (1:1 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, acetone- d_6) δ 7.70 (d, 1H, $J = 9.0$ Hz), 6.54 (dd, 1H, $J = 2.3, 8.8$ Hz), 6.37 (d, 1H, $J = 2.5$ Hz), 4.51 (t, 2H, $J = 6.3$ Hz), 2.67 (t, 2H, $J = 6.5$ Hz); $^{13}\text{C NMR}$ (125 MHz, acetone- d_6) δ 190.2, 164.9, 164.7, 129.6, 115.6, 111.0, 103.4, 68.1, 38.0; HRMS (FAB $^+$) m/z calcd for $\text{C}_9\text{H}_9\text{O}_3$ $[\text{M}]^+$ 165.0552, found 165.0553.

Compound **5**, 3-bromo-7-hydroxychroman-4-one

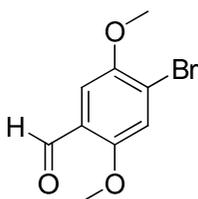


The compound **4** (1.00 mmol, 1.0 equiv.) was dissolved in a mixed solution of ethyl acetate/chloroform/methanol (5:5:1). Copper (II) bromide (CuBr_2 , 2.1 equiv.) was added in this solution and the reaction mixture was heated to reflux (70 °C) for 4 h. After the completion of the reaction monitored by TLC, the reaction mixture was filter and concentrated *in vacuo*.

The resulting residue was dissolved in ethyl acetate and washed 3 times with brine. The combined organic layer was dried over anhydrous MgSO_4 , then filtered, and concentrated under reduced pressure. The desired product **5** was obtained by recrystallization with EtOAc/hexane (1:5).

Yield: 67%, $R_f = 0.48$ (1:1 = EtOAc:hexane, v/v); ^1H NMR (500 MHz, acetone- d_6) δ 7.76 (d, 2H, $J = 8.6$ Hz), 6.65 (dd, 1H, $J = 2.3, 8.8$ Hz), 6.46 (d, 2H, $J = 2.0$ Hz), 4.75–4.81 (m, 2H), 4.62 (dd, 1H, $J = 3.5, 13.0$ Hz); ^{13}C NMR (125 MHz, acetone- d_6) δ 184.2, 165.8, 163.6, 130.8, 112.2, 103.4, 72.5, 46.9, 46.9; HRMS (FAB $^+$) m/z calcd for $\text{C}_9\text{H}_7\text{BrO}_3$ $[\text{M}]^+$ 241.9579, found 241.9576.

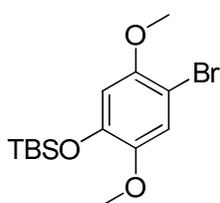
Compound **6**, 4-bromo-2,5-dimethoxybenzaldehyde



To a solution of 2,5-dimethoxybenzaldehyde (6.00 mmol) in glacial acetic acid (10 mL) was added bromine (1.1 equiv.). The reaction mixture was stirred at room temperature for 12 h. Dilution with ice water was poured into reaction mixture, and then the yellow precipitate was collected by filtration. The mono-brominated desired product **6** was obtained by recrystallization with ethanol.

Yield: 62%, $R_f = 0.56$ (1:6 = EtOAc:hexane, v/v); ^1H NMR (500 MHz, CDCl_3) δ 10.40 (s, 1H), 7.34 (s, 1H), 7.25 (s, 1H), 3.90 (s, 3H), 3.90 (s, 3H); ^{13}C NMR (125 MHz, CDCl_3) δ 188.8, 156.4, 150.7, 124.4, 120.5, 117.9, 109.9, 56.9, 56.6; HRMS (FAB $^+$) m/z calcd for $\text{C}_9\text{H}_9\text{BrO}_3$ $[\text{M}]^+$ 244.9813, found 244.9819.

Compound **7**, (4-bromo-2,5-dimethoxyphenoxy)(tert-butyl) dimethylsilane



To a solution of compound 6 (1.530 mmol, 1 equiv.) and *meta*-chloroperbenzoic acid (1.5 equiv.) in DCM (5 mL) was stirred at room temperature for 3 h. The reaction mixture was diluted with ethyl acetate and saturated NaHCO₃ aqueous solution, and extracted three times with ethyl acetate. The combined organic layer was washed with brine once, then dried over anhydrous MgSO₄, filtered, and concentrated under reduced pressure. The evaporated residue was dissolved in methanol (5 mL), and then addition of NaOH (1 equiv.) for ester hydrolysis. The reaction mixture was stirred for 3 h. After the completion of the reaction monitored by TLC, the resulting mixture was diluted with ethyl acetate, and washed with saturated NH₄Cl aqueous solution and brine. The combined organic layer was dried over anhydrous MgSO₄, then filtered, and concentrated under reduced pressure. The resulting product was used directly for the next reaction without further purification. The remained evaporated residue and imidazole (1.5 equiv.) was stirred in anhydrous DCM (20 mL) for 15 min at room temperature. To the solution was added TBSCl (1.1 equiv.) and the reaction mixture was stirred for 1 h at room temperature. Then the solution was extracted with excess ethyl acetate and washed with saturated NH₄Cl aqueous solution and brine. The organic layer was dried over anhydrous MgSO₄ and filtered. Then, the filtrate was condensed under reduced pressure. The desired product 7 was obtained by the purification using silica-gel flash column chromatography (EtOAc:hexane = 1:50).

Yield: 95% (3-step yield), R_f = 0.71 (1:6 = EtOAc:hexane, v/v); ¹H NMR (500 MHz, CDCl₃) δ 7.02 (s, 1H), 6.50 (s, 1H), 3.81 (s, 3H), 3.75 (s, 3H), 0.99 (s,

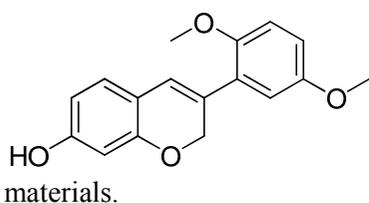
9H), 0.16 (s, 6H); ^{13}C NMR (125 MHz, CDCl_3) δ 150.3, 145.7, 145.3, 117.4, 106.9, 102.0, 56.9, 56.5, 25.8, 18.6, -4.5; HRMS (FAB $^+$) m/z calcd for $\text{C}_{14}\text{H}_{23}\text{BrO}_3\text{Si}$ $[\text{M}]^+$ 346.0600, found 346.0598.

General procedure for synthesis of compounds 8–19 via Suzuki Coupling

Reaction

Compound **2a** or **2b** (0.100 mmol, 1 equiv.), one of boronic acids/boronic esters (2 equiv.), $\text{Pd}(\text{PPh}_3)_4$ (5 mol %), Na_2CO_3 (3 equiv.) were suspended in a solvent mixture of Toluene/EtOH/ H_2O (2:1:1, total volume of 4 mL). The reaction mixture was stirred at 80 °C for 2 h ~ 8 h and the completion of the reaction was monitored by TLC. When the starting material **2a** or **2b** was all consumed, the resulting mixture was diluted with ethyl acetate and washed with brine. The combined organic layer was dried over anhydrous MgSO_4 , then filtered, and concentrated under reduced pressure. After a short silica-gel filtration, the resulting product was used directly for the next reaction without further purification. The solution of previous prepared compound in HF/Pyridine/THF (1:1:18, volumetric ratio, total volume of 3 mL) was stirred for 5 h at room temperature in a plastic vessel. After the reaction completion, Additional fluoride source was quenched with excess TMSOEt (6 mL). The mixture was evaporated under reduced pressure, the resultant was purified with silica gel flash column chromatography (with EtOAc:hexane) to provide the desired products, compounds **8–19**.

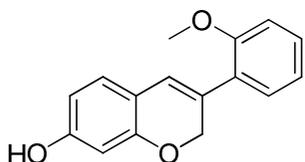
Compound **8**, 3-(2,5-dimethoxyphenyl)-2H-chromen-7-ol



Compound **8** was prepared by the general procedure with **2a** and 2,5-dimethoxyphenylboronic acid as the starting materials.

Yield: 69% (2-step yield), $R_f = 0.49$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 6.95 (d, 1H, $J = 8.0$ Hz), 6.88 (s, 1H), 6.82 (d, 1H, $J = 2.0$ Hz), 6.59 (s, 1H), 6.38–6.41 (m, 2H), 5.01 (d, 2H, $J = 1.5$ Hz), 3.80 (s, 3H), 3.78 (s, 3H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 156.7, 155.2, 153.9, 151.7, 129.5, 128.8, 128.0, 122.6, 117.1, 114.7, 113.7, 112.2, 108.7, 103.2, 68.3, 56.2, 56.0; HRMS (FAB $^+$) m/z calcd for $\text{C}_{17}\text{H}_{16}\text{O}_4$ $[\text{M}]^+$ 284.1049, found 284.1052.

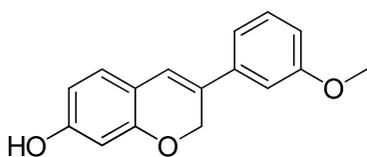
Compound **9**, 3-(2-methoxyphenyl)-2H-chromen-7-ol



Compound **9** was prepared by the general procedure with **2a** and 2-methoxyphenylboronic acid as the starting materials.

Yield: 70% (2-step yield), $R_f = 0.25$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.25–7.32 (m, 2H), 6.87–6.99 (m, 3H), 6.57 (s, 1H), 6.39–6.42 (m, 2H), 5.01 (d, 2H, $J = 1.2$ Hz), 3.82 (s, 3H); $^{13}\text{C NMR}$ (100 MHz, CDCl_3) δ 157.3, 156.7, 155.0, 129.6, 129.1, 128.9, 128.0, 127.9, 122.4, 121.1, 117.1, 111.0, 108.7, 103.2, 68.4, 55.5; HRMS (FAB $^+$) m/z calcd for $\text{C}_{16}\text{H}_{14}\text{O}_3$ $[\text{M}]^+$ 254.0943, found 254.0947.

Compound **10**, 3-(3-methoxyphenyl)-2H-chromen-7-ol

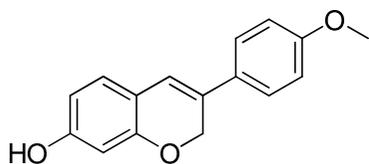


materials.

Compound **10** was prepared by the general procedure with **2a** and 3-methoxyphenylboronic acid as the starting

materials.
Yield: 55% (2-step yield), $R_f = 0.24$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.25–7.32 (m, 1H), 6.92–7.01 (m, 3H), 6.85 (dd, 1H, $J = 8.2, 2.6$ Hz), 6.76 (s, 1H), 6.37–6.42 (m, 2H), 5.12 (d, 2H, $J = 1.2$ Hz), 5.00 (br. s, 1H) 3.84 (s, 3H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 160.0, 156.7, 154.8, 138.5, 129.8, 128.7, 128.2, 120.3, 117.3, 116.5, 113.0, 110.6, 108.8, 103.2, 67.4, 55.5; HRMS (FAB⁺) m/z calcd for $\text{C}_{16}\text{H}_{14}\text{O}_3$ [M]⁺ 254.0943, found 254.0945.

Compound **11**, 3-(4-methoxyphenyl)-2H-chromen-7-ol

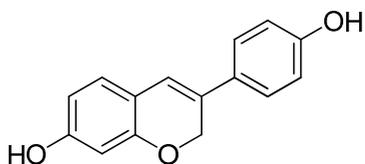


materials.

Compound **11** was prepared by the general procedure with **2a** and 4-methoxyphenylboronic acid as the starting

materials.
Yield: 62% (2-step yield), $R_f = 0.24$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, acetone- d_6) δ 7.46 (dd, 2H, $J = 3.8, 8.3$ Hz), 6.94–6.99 (m, 3H), 6.81 (s, 1H), 6.42 (dd, 2H, $J = 2.5, 8.0$ Hz), 6.33 (d, 1H, $J = 2.5$ Hz), 5.08 (d, 2H, $J = 4.5$ Hz), 3.81 (s, 3H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 160.3, 155.4, 130.3, 128.7, 126.6, 118.9, 116.5, 115.0, 109.5, 103.5, 67.7, 55.7; HRMS (FAB⁺) m/z calcd for $\text{C}_{16}\text{H}_{14}\text{O}_3$ [M]⁺ 254.0943, found 254.0948.

Compound **12**, 3-(4-hydroxyphenyl)-2H-chromen-7-ol



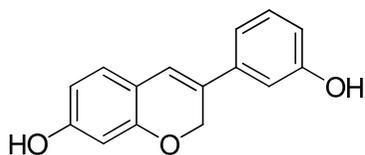
Synthetic procedure of used boronic ester for compound **12** was as follows: To a solution of

4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenol (1.00 mmol, 1 equiv.) and imidazole (1.5 equiv.) in DMF (3 mL) was added TBSCl (1.1 equiv.) and the reaction mixture was stirred for 5 h at room temperature. Then the solution was extracted with excess ethyl acetate and washed with saturated NH_4Cl aqueous solution and brine. The organic layer was dried over anhydrous MgSO_4 and filtered. Then, the filtrate was condensed under reduced pressure. The desired boronic ester *tert*-butyldimethyl(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)silane was obtained by the purification using silica-gel flash column chromatography (EtOAc:hexane = 1:50). Yield: 84%, R_f = 0.48 (1:30 = EtOAc:hexane, v/v); ^1H NMR (400 MHz, CDCl_3) δ 7.69 (d, 2H, J = 8.4 Hz), 6.83 (d, 2H, J = 8.4 Hz), 1.33 (s, 12H), 0.98 (s, 9H), 0.20 (s, 6H); ^{13}C NMR (100 MHz, CDCl_3) δ 158.7, 136.6, 119.8, 83.7, 60.5, 25.8, 25.0, 18.4, -4.3.

Compound **12** was prepared by the general procedure with **2a** and prepared *tert*-butyldimethyl(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)silane as the starting materials.

Yield: 57% (2-step yield), R_f = 0.42 (1:1 = EtOAc:hexane, v/v); ^1H NMR (500 MHz, CD_3OD) δ 7.29 (d, 2H, J = 8.5 Hz), 6.90 (d, 1H, J = 8.0 Hz), 6.77–6.81 (m, 2H), 6.66 (s, 1H), 6.26–6.35 (m, 1H), 6.26 (d, 1H, J = 2.0 Hz), 5.02 (s, 2H); ^{13}C NMR (125 MHz, CD_3OD) δ 159.2, 158.2, 155.7, 129.4, 128.6, 127.5, 127.5, 126.8, 118.5, 117.0, 116.5, 109.6, 103.5, 68.2; HRMS (FAB $^+$) m/z calcd for $\text{C}_{15}\text{H}_{12}\text{O}_3$ [M] $^+$ 240.0786, found 240.0789.

Compound **13**, 3-(3-hydroxyphenyl)-2H-chromen-7-ol

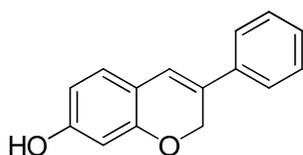


materials.

Compound **13** was prepared by the general procedure with **2a** and 3-hydroxyphenylboronic acid as the starting

Yield: 58% (2-step yield), $R_f = 0.46$ (1:1 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, DMSO- d_6) δ 7.17 (t, 1H, $J = 8.0$ Hz), 6.98 (d, 1H, $J = 6.4$ Hz), 6.92 (d, 1H, $J = 6.4$ Hz), 6.84–6.86 (m, 2H), 6.34 (dd, 1H, $J = 1.6, 6.8$ Hz), 6.25 (s, 1H), 5.03 (s, 2H); $^{13}\text{C NMR}$ (75 MHz, DMSO- d_6) δ 158.8, 157.7, 154.2, 137.7, 129.6, 128.1, 127.1, 119.5, 115.1, 114.6, 114.4, 111.0, 108.8, 102.4, 66.3; HRMS (FAB $^+$) m/z calcd for $\text{C}_{15}\text{H}_{12}\text{O}_3$ $[\text{M}]^+$ 240.0786, found 240.0788.

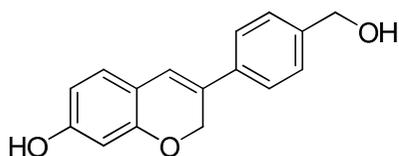
Compound **14**, 3-phenyl-2H-chromen-7-ol



Compound **14** was prepared by the general procedure with **2a** and phenylboronic acid as the starting materials.

Yield: 51% (2-step yield), $R_f = 0.71$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, CD_2Cl_2) δ 7.27–7.44 (m, 5H), 6.98 (d, 1H, $J = 8.5$ Hz), 6.80 (s, 1H), 6.41 (dd, 1H, $J = 2.5, 8.0$ Hz), 6.36 (d, 1H, $J = 1.5$ Hz), 5.14 (s, 2H); $^{13}\text{C NMR}$ (125 MHz, CD_2Cl_2) δ 157.2, 137.2, 129.1, 128.4, 128.0, 124.8, 120.0, 126.7, 108.9, 103.2, 67.6; HRMS (FAB $^-$) m/z calcd for $\text{C}_{15}\text{H}_{12}\text{O}_3$ $[\text{M}]^-$ 223.0759, found 223.0766.

Compound **15**, 3-(4-(hydroxymethyl)phenyl)-2H-chromen-7-ol

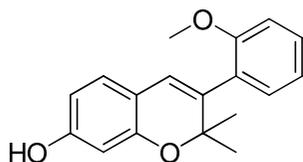


starting materials.

Compound **15** was prepared by the general procedure with **2a** and 4-(hydroxymethyl)phenylboronic acid as the

Yield: 57% (2-step yield), $R_f = 0.41$ (1:10 = MeOH:DCM, v/v); $^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 7.46 (d, 2H, $J = 8.4$ Hz), 7.32 (d, 2H, $J = 8.4$ Hz), 6.98 (d, 1H, $J = 8.0$ Hz), 6.94 (s, 1H), 6.35 (dd, 2H, $J = 2.0, 8.0$ Hz), 6.26 (d, 1H, $J = 2.0$ Hz), 5.08 (d, 1H, $J = 1.2$ Hz), 4.50 (s, 2H); $^{13}\text{C NMR}$ (100 MHz, DMSO- d_6) δ 158.7, 154.1, 141.9, 134.7, 128.0, 127.0, 126.7, 124.0, 119.1, 114.5, 108.8, 102.4, 66.3, 62.6; HRMS (FAB $^+$) m/z calcd for $\text{C}_{16}\text{H}_{14}\text{O}_3$ $[\text{M}]^+$ 254.0943, found 254.0941.

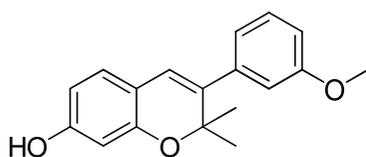
Compound **16**, 3-(2-methoxyphenyl)-2,2-dimethyl-2H-chromen-7-ol



Compound **16** was prepared by the general procedure with **2b** and 2-methoxyphenylboronic acid as the starting materials.

Yield: 71% (2-step yield), $R_f = 0.22$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.30 (t, 1H, $J = 7.8$ Hz), 7.14 (d, 1H, $J = 6.0$ Hz), 6.89–6.96 (m, 3H), 6.39 (br. s, 1H), 6.36 (br. s, 1H), 6.21 (s, 1H), 3.81 (s, 3H), 1.26 (s, 6H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 157.4, 156.5, 154.2, 137.0, 131.7, 129.1, 129.0, 127.4, 122.7, 120.5, 116.8, 110.7, 108.1, 103.9, 55.3, 26.2; HRMS (FAB $^+$) m/z calcd for $\text{C}_{18}\text{H}_{18}\text{O}_3$ $[\text{M}]^+$ 282.1256, found 282.1267.

Compound **17**, 3-(3-methoxyphenyl)-2,2-dimethyl-2H-chromen-7-ol

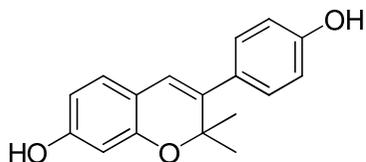


materials.

Compound **17** was prepared by the general procedure with **2b** and 3-methoxyphenylboronic acid as the starting

Yield: 75% (2-step yield), $R_f = 0.22$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.23–7.28 (m, 1H), 6.84–6.93 (m, 4H), 6.38 (br. s, 2H), 6.29 (s, 1H), 3.83 (s, 3H), 1.52 (s, 6H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 159.4, 156.7, 154.0, 141.2, 139.2, 129.2, 127.5, 122.1, 120.8, 116.5, 114.4, 112.6, 108.3, 103.9, 79.2, 55.4, 27.2; HRMS (FAB $^+$) m/z calcd for $\text{C}_{18}\text{H}_{18}\text{O}_3$ $[\text{M}]^+$ 282.1256, found 282.1250.

Compound **18**, 3-(4-hydroxyphenyl)-2,2-dimethyl-2H-chromen-7-ol



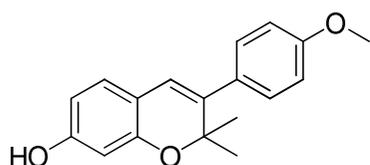
compound **12**) as the starting materials.

Compound **18** was prepared by the general procedure with **2b** and prepared *tert*-

butyldimethyl(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy) silane (Same boronic ester with synthesis of compound **12**) as the starting materials.

Yield: 88% (2-step yield), $R_f = 0.54$ (1:1 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.13 (dd, 2H, $J = 2.0, 6.8$ Hz), 6.86 (d, 1H, $J = 8.4$ Hz), 6.75 (dd, 2H, $J = 2.2, 6.6$ Hz), 6.32 (dd, 1H, $J = 2.4, 7.6$ Hz), 6.25 (d, 1H, $J = 2.4$ Hz), 6.18 (s, 1H), 1.45 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 159.4, 157.9, 154.9, 140.0, 132.4, 130.4, 128.1, 122.2, 117.0, 115.9, 109.2, 104.3, 79.9, 27.4; HRMS (FAB $^+$) m/z calcd for $\text{C}_{17}\text{H}_{16}\text{O}_3$ $[\text{M}]^+$ 268.1099, found 268.1096.

Compound **19**, 3-(4-methoxyphenyl)-2,2-dimethyl-2H-chromen-7-ol



materials

Compound **19** was prepared by the general procedure with **2b** and 4-methoxyphenylboronic acid as the starting

Yield: 79% (2-step yield), $R_f = 0.22$ (1:10 = EtOAc:hexane, v/v); $^1\text{H NMR}$ (400 MHz, CD_3OD) δ 7.22 (dd, 2H, $J = 2.0, 6.8$ Hz), 6.87 (dd, 2H, $J = 1.8, 6.6$ Hz), 6.85 (s, 1H), 6.32 (dd, 1H, $J = 2.4, 8.0$ Hz), 6.26 (d, 1H, $J = 2.0$ Hz), 6.20 (s, 1H), 4.89 (br. s, 1H), 3.78 (s, 3H), 1.45 (s, 6H); $^{13}\text{C NMR}$ (100 MHz, CD_3OD) δ 160.4, 159.5, 154.9, 139.6, 133.5, 130.4, 128.2, 122.5, 116.8, 114.5, 109.2, 104.3, 79.8, 55.7, 27.4; HRMS (FAB $^+$) m/z calcd for $\text{C}_{18}\text{H}_{18}\text{O}_3$ $[\text{M}]^+$ 282.1256, found 282.1260.

3.1.6. Supporting Information

Amino acids in 2 Å range from Eryvarin H

CYS269	ALA272	ASP273	GLU275	TRP305	LEU309	TYR326
LEU342	LEU345	HIS434	PHE435	LEU440	GLU441	

Amino acids in 2 Å range from GSK5182

CYS269	ALA272		GLU275		LEU309	TYR326
LEU342	LEU345	HIS434	PHE435	LEU440	GLU441	

Amino acids in 4 Å range from Eryvarin H

LEU265	LEU268	CYS269	ASP270	LEU271	ALA272	ASP273
ARG274	GLU275	LEU276	TRP305	MET306	LEU309	ILE310
VAL313	ARG316	TYR326	MET332	LEU342	LEU345	ASN346
ILE349	ALA431	HIS434	PHE435	ILE438	LEU440	GLU441
MET446	LEU449					

Amino acids in 4 Å range from GSK5182

LEU265	LEU268	CYS269		LEU271	ALA272	ASP273
	GLU275	LEU276	TRP305	MET306	LEU309	ILE310
VAL313	ARG316	TYR326	MET332	LEU342	LEU345	ASN346
ILE349	ALA431	HIS434	PHE435	ILE438	LEU440	GLU441
MET446	LEU449					

Supplementary Figure S3.1.1. List of interacting amino acids with Eryvarin H and GSK5182, in the range of 2 Å and 4 Å. These diagrams were obtained from Discovery Studio Version 1.5 based on docked ligands to ERRγ.

3.1.7. References and Notes

† Portions of this chapter have been previously reported, see: Koo, J. Y.;

- Oh, S.; Cho, S.-L.; Koh, M.; Oh, W.-K.; Choi, H.-S.; Park, S. B. *Org. Biomol. Chem.* **2013**, *11*, 5782–5786.
- (1) Horard, B.; Vanacker, J.-M. *J. Mol. Endocrinol.* **2003**, *31*, 349–357.
 - (2) Shi, Y. *Drug Discov. Today* **2007**, *12*, 440–445.
 - (3) Moore, J. T.; Collins, J. L.; Pearce, K. H. *ChemMedChem* **2006**, *1*, 504–523.
 - (4) Schreiber, S. L. *Bioorg. Med. Chem.* **1998**, *6*, 1127–1152.
 - (5) Giguère, V. *Trends Endocrin. Met.* **2002**, *13*, 220–225.
 - (6) Ariazi, E. A.; Jordan, V. C. *Curr. Top. Med. Chem.* **2006**, *6*, 203–215.
 - (7) (a) Xie, Y.-B.; Park, J.-H.; Kim, D.-K.; Hwang, J. H.; Oh, S.; Park, S. B.; Shong, M.; Lee, I.-K.; Choi, H.-S. *J. Biol. Chem.* **2009**, *284*, 28762–28774. (b) Kim, D.-K.; Kim, J. R.; Koh, M.; Kim, Y. D.; Lee, J.-M.; Chanda, D.; Park, S. B.; Min, J.-J.; Lee, C.-H.; Park, T.-S.; Choi, H.-S. *J. Biol. Chem.* **2011**, *286*, 38035–38042. (c) Kim, D.-K.; Ryu, D.; Koh, M.; Lee, M.-W.; Lim, D.; Kim, M.-J.; Kim, Y.-H.; Cho, W.-J.; Lee, C.-H.; Park, S. B.; Koo, S.-H.; Choi, H.-S. *J. Biol. Chem.* **2012**, *287*, 21628–21639. (d) Lee, J. H.; Kim, E.-J.; Kim, D.-K.; Lee, J.-M.; Park, S. B.; Lee, I.-K.; Harris, R. A.; Lee, M.-O.; Choi, H.-S. *PLoS ONE* **2012**, *7*, e46324. (e) Kim, D.-K.; Gang, G.-T.; Ryu, D.; Koh, M.; Kim, Y.-N.; Kim, S. S.; Park, J.; Kim, Y.-H.; Sim, T.; Lee, I.-K.; Choi, C. S.; Park, S. B.; Lee, C.-H.; Koo, S.-H.; Choi, H.-S. *Diabetes* **2013**, *62*, 3093–3102.
 - (8) (a) Giguère, V. *Endocr. Rev.* **2008**, *29*, 677–696. (b) Tiraby, C.; Hazen, B. C.; Gantner, M. L.; Kralli, A. *Cancer. Res.* **2011**, *71*, 2518–2528.

- (c) Ariazi, E. A.; Clark, G. M.; Mertz, J. E. *Cancer. Res.* **2002**, *62*, 6510–6518. (d) Yu, S.; Wang, X.; Ng, C.-F.; Chen, S.; Chan, F. L. *Cancer. Res.* **2007**, *67*, 4904–4914.
- (9) Heard, D. J.; Norby, P. L.; Holloway, J.; Vissing, H. *Mol. Endocrine.* **2000**, *14*, 382–392.
- (10) (a) Kim, Y.; Koh, M.; Kim, D.-K.; Choi, H.-S.; Park, S. B. *J. Comb. Chem.* **2009**, *11*, 928–937. (b) Koh, M.; Park, S. B. *Mol. Divers.* **2011**, *15*, 69–81.
- (11) (a) Greschik, H.; Flaig, R.; Renaud, J.-P.; Moras, D. *J. Biol. Chem.* **2004**, *279*, 33639–33649. (b) Klinge, C. M.; Jernigan, S. C.; Risinger, K. E. *Endocrinology* **2002**, *143*, 853–867. (c) Coward, P.; Lee, D.; Hull, M. V.; Lehmann, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8880–8884. (d) Yu, D. D.; Forman, B. M. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1311–1313. (e) Takayanagi, S.; Tokunaga, T.; Liu, X.; Okada, H.; Matsushima, A.; Shimohigashi, Y. *Toxicol. Lett.* **2006**, *167*, 95–105. (f) Wang, J.; Fang, F.; Huang, Z.; Wang, Y.; Wong, C. *FEBS Lett.* **2009**, *583*, 643–647. (g) Wang, Y.; Fang, F.; Wong, C.-W. *Biochem. Pharmacol.* **2010**, *80*, 80–85. (h) Hirvonen, J.; Rajalin, A.-M.; Wohlfahrt, G.; Adlercreutz, H.; Wähälä, K.; Aarnisalo, P. *J. Steroid Biochem.* **2011**, *123*, 46–57.
- (12) Chao, E. Y. H.; Collins, J. L.; Gaillard, S.; Miller, A. B.; Wang, L.; Orband-Miller, L. A.; Nolte, R. T.; McDonnell, D. P.; Willson, T. M.; Zuercher, W. J. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 821–824.
- (13) Tanaka, H.; Hirata, M.; Etoh, H.; Sako, M.; Sato, M.; Murata, J.;

- Murata, H.; Darnaedi, D.; Fukai, T. *Heterocycles* **2003**, *60*, 2767–2773.
- (14) Yenesew, A.; Derese, S.; Irungu, B.; Midiwo, J. O.; Waters, N. C.; Liyala, P.; Akala, H.; Heydenreich, M.; Peter, M. G. *Planta Med.* **2003**, *69*, 658–661.
- (15) Yenesew, A.; Twinomuhwezi, H.; Kiremire, B. T.; Mbugua, M. N.; Gitu, P. M.; Heydenreich, M.; Peter, M. G. *Bull. Chem. Soc. Ethiop.* **2009**, *23*, 205–210.
- (16) Koch, K.; Biggers, M. S. *J. Org. Chem.* **1994**, *59*, 1216–1218.
- (17) Sardessai, M. S.; Abramson, H. N. *Org. Prep. Proced. Int.* **1991**, *23*, 419–424.
- (18) Ko, S. K.; Jang, H. J.; Kim, E.; Park, S. B. *Chem. Commun.*, **2006**, 2962–2964.
- (19) Sambrook, J.; Russel, D. W. *Molecular Cloning: A Laboratory Manual*, CSHL Press, New York, 3rd ed., **2001**, vol. 3, pp. 16.21–16.26.
- (20) Sambrook, J.; Russel, D. W. *Molecular Cloning: A Laboratory Manual*, CSHL Press, New York, 3rd ed., **2001**, vol. 3, pp. 17.48–17.51.
- (21) Gronemeyer, H. ; Gustafsson, J.-A.; Laudet, V. *Nat. Rev. Drug Discov.* **2004**, *3*, 950–964.

3.2. Treatment of Sepsis Pathogenesis with High Mobility Group Box Protein 1 (HMGB1)-regulating Anti-Inflammatory Agents

3.2.1. Abstract

Sepsis is one of the major causes of death worldwide when associated with multiple organ failure. However, there is a critical lack of adequate sepsis therapies because of its diverse patterns of pathogenesis. The pro-inflammatory cytokine cascade mediates sepsis pathogenesis, and high mobility group box proteins (HMGBs) play an important role as late-stage cytokines. We previously reported the small-molecule modulator, inflachromene (**1d**), which inhibits the release of HMGBs and, thereby, reduces the production of pro-inflammatory cytokines. In this context, we intraperitoneally administered **1d** to a cecal ligation and puncture (CLP)-induced mouse model of sepsis and confirmed that it successfully ameliorated sepsis pathogenesis. Based on a structure-activity relationship study, we discovered new candidate compounds, **2j** and **2l**, with improved therapeutic efficacy *in vivo*. Therefore, our study clearly demonstrates that the regulation of HMGB1 release using small molecules is a promising strategy for the treatment of sepsis.

3.2.2. Introduction

Sepsis, a systemic inflammatory response caused by infectious processes, is

one of the major causes of death worldwide.¹⁻³ External pathogenic molecules known as pathogen-associated molecular patterns (PAMPs) are recognized by pattern recognition receptors (PRRs) in innate immune cells. The specific recognition of PAMP by the PRRs activates immune signalling, resulting in the production of various pro-inflammatory cytokines that mediate the inflammatory responses.^{4,5} Sepsis occurs when immune responses are over-activated, including intractable inflammatory responses associated with imbalanced cytokine production. In addition, the subsequently unmanaged pro-inflammatory cytokine cascade results in whole body shock.^{1,6,7} When the pathophysiology of sepsis is associated with hypoperfusion, hypotension, and organ dysfunction, it is termed as severe sepsis,¹ which occurs in approximately 10% of all intensive care unit patients in the US.⁷⁻⁹ Although numerous therapeutic approaches to sepsis have been advanced, the mortality rate of severe sepsis is nearly 20%.¹⁰ Furthermore, the annual hospital healthcare cost for patients with severe sepsis in the US is the highest among all diseases and was around \$20 billion.¹¹ The biggest obstacle to the discovery of therapeutics for sepsis is its diverse etiology among patients.¹² The heterogeneous patterns of sepsis pathogenesis depend on the pathogenic organisms and sites of infection.^{7,12} The diverse characteristics of sepsis pathogenesis have prompted researchers to study the molecular mechanism of sepsis progression based on systemic inflammatory responses.

The emerging role of various pro-inflammatory cytokines in the progression of sepsis has been examined by a series of scientific studies.¹³⁻¹⁵ Furthermore, tumor necrosis factor (TNF)- α and interleukin-(IL)-1 β are two major pro-

inflammatory cytokines involved in sepsis and their secretion is regulated by positive feedback during systemic inflammatory responses.^{15,16} When the secreted cytokines circulate in the whole body, considerable amounts of cytokines are produced, resulting in septic shock and secondary multiple organ dysfunction.¹⁶ Therefore, therapeutic approaches to antagonize the production of these cytokines have been used for the treatment of sepsis.^{17,18} For example, an anti-TNF- α antibody underwent large-scale clinical trials, but was discontinued because it lacked efficacy.^{7,19,20} Literature reports suggest these failures might have been caused by two reasons. First, because the production of TNF- α and IL-1 β peaks at the early stage of systemic inflammation, it is difficult to reverse the enhanced cytokine production without inhibiting the early stage of pathogenesis. Second, the continuous inflammatory stimuli unexpectedly restored the TNF- α levels after treatment with the antagonizing anti-TNF- α antibody.¹⁹⁻²¹

On the other hand, high mobility group box protein 1 (HMGB1) has been identified as a late-stage mediator of inflammatory responses.²² HMGB1 is a DNA-binding protein that is associated with the nuclear genomic DNA. When damaged cells undergo necrotic death, HMGB1 is released into the extracellular milieu, alerting adjacent cells to the damage. Interestingly, in immune cells such as macrophages, HMGB1 is actively secreted as a pro-inflammatory cytokine during systemic inflammatory responses.²³⁻²⁶ Previous reports have indicated that secreted HMGB1 plays an important role in regulating the production of other pro-inflammatory cytokines.²⁷ In this context, it has been reported that the inhibition of HMGB1 secretion successfully

ameliorated the pathogenesis of sepsis in an *in vivo* cecal ligation and puncture (CLP) mouse model—the gold standard animal model in the field of sepsis study.^{28,29} Briefly, the CLP mouse model is established by performing a surgical procedure to the cecum, which is the organ where the resident enteric microbiome exists. Ligating the cecum and puncturing the end with a needle causes endotoxemia by inducing polymicrobial contamination in the abdominal cavity towards the circulatory system. This action subsequently initiates a systemic inflammatory response, namely sepsis.

Regarding the role of HMGB1 in sepsis, Tracey and coworkers reported that the anti-HMGB1 antibody abrogated the lethality of the CLP-induced mouse model by inhibiting the release of HMGB1.³⁰ Endogenous small molecules such as cholinergic agonists have also inhibited HMGB1 secretion, thereby exhibiting the therapeutic effects in sepsis.³¹ Furthermore, a known autophagy modulator, (-)-epigallocatechin-3-gallate (EGCG), has been studied for the prevention of lipopolysaccharide (LPS)-induced HMGB1 release by triggering the degradation of cytosolic HMGB1 thereby boosting the autophagic process.^{32–34}

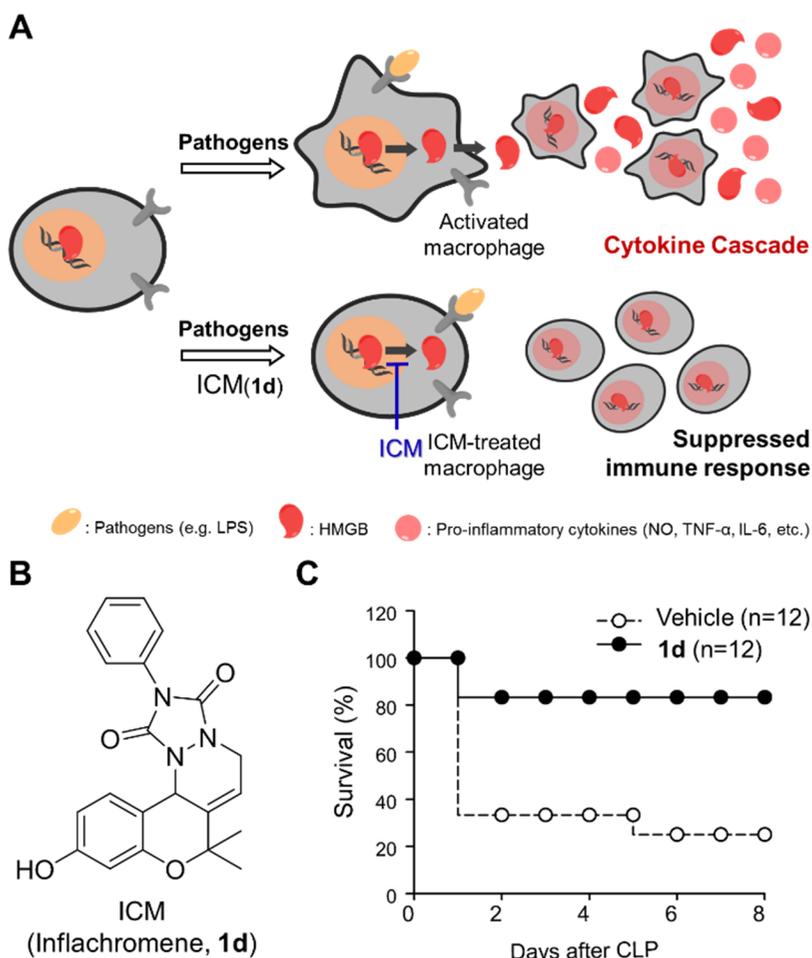


Figure 3.2.1. Structure and *in vivo* therapeutic effect of inflachromene (ICM, **1d) in cecal ligation and puncture (CLP) model of sepsis.** (A) During sepsis the extracellular secretion of high-mobility group box protein 1 (HMGB1) occurs in activated macrophages. The small-molecule ICM (**1d**) inhibits the release of HMGB1, thereby suppressing immune responses. (B) Chemical structure of ICM. (C) Survival rate of vehicle- and ICM-treated CLP mouse model. ICM was administered intraperitoneally at a daily dose of 10 mg/kg for 9 days.

Previously, we reported the discovery of a novel small molecule, inflachromene (ICM, **1d**, Figure 3.2.1B), that inhibits the activation of BV-2 microglia-like cells, via screening a pDOS-based drug-like compound library.^{35,36} Based on our efforts on target identification using the fluorescence difference in two-dimensional gel electrophoresis (FITGE) method,³⁷ we revealed HMGB1 and HMGB2 as the target proteins of ICM, and the subsequent biochemical and biophysical studies confirmed that ICM inhibits the secretion of HMGB1 and HMGB2 in microglia via the modulation of their post-translational modifications.³⁸ Based on the fact that microglia in the central nervous system is a cell type sharing the same ontology with peripheral macrophages,^{39,40} we hypothesized that ICM can inhibit the secretion of HMGB1 and HMGB2 in macrophages, reduce the production of additional pro-inflammatory cytokines, and thereby exhibit its therapeutic effect on the systemic inflammatory response. To test our hypothesis, we examined whether *in vivo* administration of ICM to CLP-induced mice inhibits sepsis pathogenesis by measuring the survival rate. Based on the initial proof-of-concept *in vivo* study, we investigated the enhancement of the anti-inflammatory efficacy of ICM in sepsis treatment by conducting a systematic structure-activity relationship study. After synthesizing approximately 30 ICM analogues, we examined their anti-inflammatory activities in Raw264.7 macrophage-like cells. Then, candidate compounds were selected based on a series of biological evaluations including the Griess assay, enzyme-linked immunosorbent assay (ELISA), microsomal stability testing, and pharmacokinetics (PK) as well as *in vivo* studies using a CLP mouse model. Finally, we discovered new candidate

compounds with enhanced therapeutic efficacy against sepsis.

3.2.3. Result and Discussion

***In vivo* therapeutic effect of ICM in CLP-induced mouse model of sepsis.**

When pathogens invade a host organism, an acute inflammatory response is initiated in the body. Macrophages, one of the key immune cells that mediate the inflammatory responses, are activated following their recognition of pathogens, which subsequently induces the secretion of various pro-inflammatory cytokines including HMGBs, TNF- α , and IL-6 (Figure 3.2.1A).⁴⁻

⁶ At the basal state, HMGBs are associated with genomic DNA within the nucleus. After activation of macrophage upon recognition of LPS, HMGBs are translocated from nucleus to cytosol and subsequently released to the extracellular milieu.⁴¹ These pro-inflammatory cytokines activate other adjacent macrophages and initiate a positive feedback mechanism of immune response.⁴² We previously reported that ICM (**1d**, Figure 3.2.1B) inhibits the translocation of HMGBs by blocking their post-translational modifications and subsequently reduces the secretion of HMGBs in microglia, which led to the suppression of neuroinflammatory response.³⁸

Therefore, we envisaged that **1d** would be potentially efficacious for the treatment of sepsis. Based on this assumption, we adopted a murine CLP model, which is considered as the gold standard model of *in vivo* sepsis.²⁸ The intraperitoneal (i.p.) injection in the CLP-induced mice of 10 mg/kg of **1d** for

9 days significantly increased the survival rate compared to that of the vehicle-treated group (Figure 3.2.1C). The measurement of the serum cytokines after euthanizing the mice revealed that IL-6 levels were highly correlated with their survival (Figure S3.2.1). Based on these data, we concluded that **1d** exhibited a potential therapeutic effect in sepsis.

Synthesis and biological evaluation of ICM (1d) analogues to discover new compounds with enhanced anti-inflammatory effects.

To improve the therapeutic efficacy of **1d**, we first comprehensively explored the structure-activity relationship of its anti-inflammatory effects. Starting from the benzopyran core structure of **1d**, we designed and synthesized a series of analogues to clarify the importance of each structural motif. To measure the inhibitory effects on immune response, Raw264.7 cells were treated with the synthesized analogues in the presence of LPS, which is widely used to activate macrophages,⁴³ because it is one of the major components of external membrane of gram-negative bacteria.⁴³⁻⁴⁵ LPS-activated macrophages produce various pro-inflammatory cytokines^{46,47} and nitric oxide (NO), which was selected as the molecular marker to monitor the activation level of macrophages. We quantitatively measured the secreted NO level in the culture medium by using the Griess assay as an efficient high-throughput screening system.^{38,48}

Our original hit compound (**1d**) inhibited the NO release by 87% in Raw264.7 cells compared to the vehicle. As shown in Table 3.2.1, the systematic modification of the R₁, R₂, R₃, and R₄ groups of the benzopyran core structure

revealed that most of R₁, R₂, and R₃ substituents (**1a–1h**) reduced the inhibition potency on NO release respect to **1d**: very unfavourable was in particular substitution of the R₂ hydroxyl group with a methoxy one. Even when both methyl groups of **1d** at the R₄ position were modified as in **1i–1j** analogues, the inhibition of cellular NO release was significantly reduced. These results indicate that the R₁–R₄ groups' combination of **1d** is essential for the anti-inflammatory effects. Similarly, the shift or removal of the double bond of the tetrahydropyridazine ring failed to induce inhibitory effects on NO release (Figure S3.2.2).

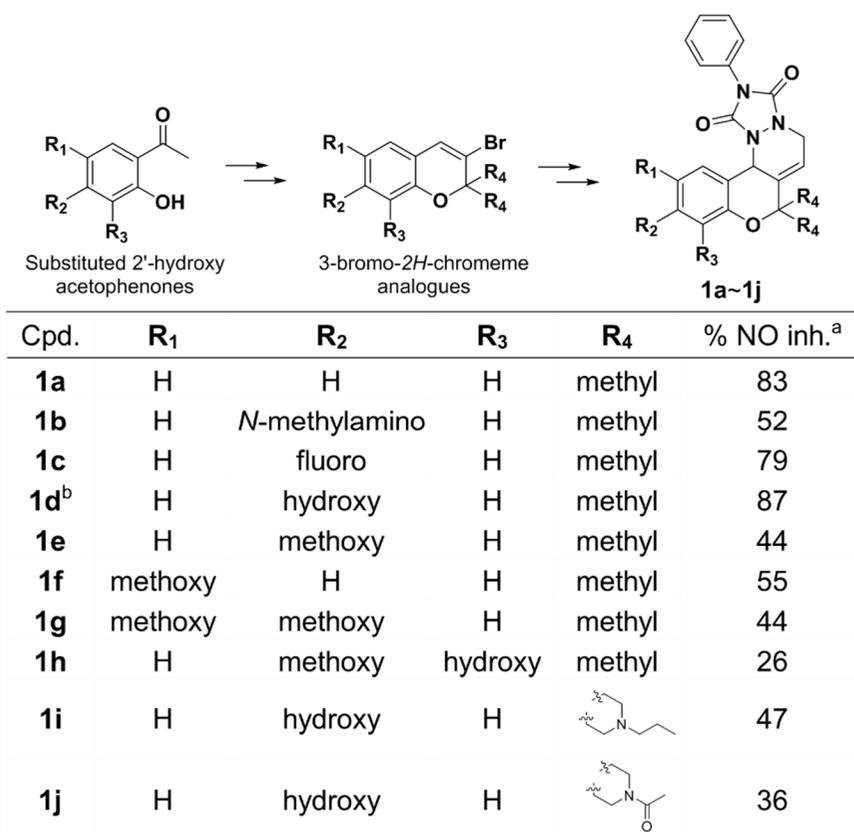
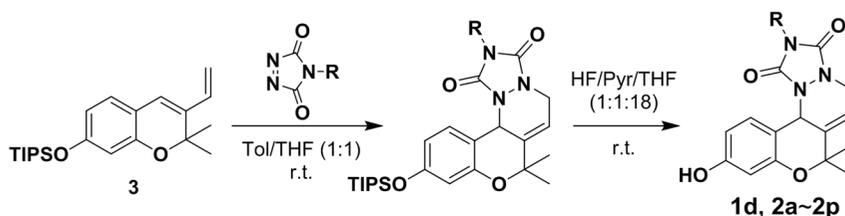


Table 3.2.1. Relationship between structural modifications in benzopyranyl core and percent (%) inhibition of released nitric oxide (NO) level. ^aPercentage inhibition of NO by 10 μ M ICM and its analogues was measured in Raw264.7 cells by using Griess assay; ^bICM, Cpd., compound; NO inh, nitric oxide inhibition.

Next, we prepared a series of **1d** analogues by introducing various *N*-substituents of the tetracyclic benzopyranyl core, indicated as the R group in Table 3.2.2, via hetero-Diels-Alder reaction of the intermediate **3** with *N*-substituted 1,2,4-triazoline-3,5-diones. First, the replacement of the phenyl group of **1d** with a methyl one (**2a**) caused the lowest inhibitory effect (45%) on NO release, being reduced also with benzyl (**2b**) and cyclohexyl (**2c**) substitutions. Then, we incorporated various substituents at the *ortho*, *meta*, and *para* positions of the original phenyl ring. The *ortho*-substituted compounds (**2d**, **2e**) exhibited little inhibitory effects on NO release. In contrast, a methoxy group at the *meta* (**2f**) and *para* (**2g**) positions enhanced the inhibition of NO release. To examine the relationship between the position of the methoxy group on the phenyl ring and the inhibition of NO release, we synthesized ICM analogues containing 3,4-(methylenedioxy)phenyl (**2k**) and 3,5-dimethoxyphenyl (**2l**) as the R groups. A series of *para*-substituted (**2g–j**, **2m–o**) analogues were also synthesized and their activity was evaluated by using the Griess assay. The screening results indicated that compounds **2g–2l** effectively inhibited the NO release compared to **1d** while **2m**, **2n**, and **2o** showed cytotoxicity at the higher concentration, which led us to discontinue their further biological evaluation. To confirm the anti-inflammatory activity of our synthesized compounds, we measured their inhibitory effect on the

production of another cytokine, TNF- α . As shown in Table 3.2.2, compounds **1d** and **2f–2l** inhibited the TNF- α release in Raw264.7 cells. The amount of TNF- α released was quantitatively measured by using ELISA. Based on these results, we selected compounds **2i**, **2j**, **2k**, and **2l** as our improved candidate compounds for further biological evaluation.



Cpd.	R	% NO inh. ^a	% viab. ^b	% TNF- α inh. ^c
1d^d	phenyl	87	105	43
2a	methyl	45	114	– ^e
2b	benzyl	70	106	– ^e
2c	cyclohexyl	63	117	– ^e
2d	2-iodophenyl	48	109	– ^e
2e	2-methoxyphenyl	66	115	– ^e
2f	3-methoxyphenyl	96	118	38
2g	4-methoxyphenyl	89	127	44
2h	4-fluorophenyl	87	109	44
2i	4-methylphenyl	94	130	45
2j	4-acetylphenyl	93	111	68
2k	3,4-(methylenedioxy)phenyl	92	121	46
2l	3,5-dimethoxyphenyl	94	121	60
2m	4-(trifluoromethyl)phenyl	85	106 ^f	–
2n	4- <i>tert</i> -butylphenyl	99	105 ^f	–
2o	4- <i>n</i> -butylphenyl	99	50 ^f	–

Table 3.2.2. Effects of structural modifications relative to the *N*-substituents of the triazolinedione ring structure of **1d on the percentage (%) inhibition of nitric oxide (NO) and tumor necrosis factor (TNF)- α release, as well as on the cell viability (%). ^aInhibition of**

NO release (%) by each compound (10 μ M) was measured by using Griess assay in Raw264.7 cells; ^bRelative cell viability (%) of each compound (10 μ M) in the presence of LPS was measured by using WST assay, and compared to vehicle; ^cInhibition of TNF- α release (%) by each compound (10 μ M) was measured by using ELISA; ^dICM; ^eExcluded by low NO inhibition (%); ^fExcluded due to the cellular toxicity observed at the higher concentration. Cpd., compound; WST, water-soluble tetrazolium; ELISA, enzyme-linked immunosorbent assay.

Inhibition of HMGB1 and IL-6 release by initial hit compounds.

For the secondary confirmation of biological activities, we first examined whether our initial hit compounds could inhibit the translocation of HMGB1. As mentioned earlier, HMGBs are translocated from the nucleus to cytosol in activated macrophage before they are released into the extracellular milieu. Therefore, we measured the cellular localization of HMGB1 upon LPS treatment in the absence and presence of our hit compounds using immunohistochemistry (TRITC-labeled secondary antibody). As shown in Figure 3.2.2A, our initial hit compounds (**2i**, **2j**, **2k**, **2l**, and **1d**) inhibited the translocation of HMGB1 from the nucleus to the cytosol induced by LPS-treatment. Because we observed a drastic decrease of IL-6 levels in the surviving CLP-induced mice treated with **1d** (Figure S3.2.1), we measured the secreted IL-6 levels in Raw264.7 cells. The western blot analysis confirmed that the secreted IL-6 levels were significantly decreased upon treatment with **1d**, **2i**, **2j**, **2k**, and **2l** in LPS-activated macrophages (Figure 3.2.2B). Interestingly, we also observed the reduction of intracellular IL-6 levels upon treatment with hit compounds, which indicates that the positive feedback of cytokine production was inhibited by our molecules. Furthermore, among the

initial hit compounds, **2j** and **2l** showed superior inhibitory efficacy on IL-6 secretion (Figure 3.2.2B) and TNF- α release (Table 3.2.2).

To confirm the improved activity of **2j** and **2l**, we performed *in vitro* binding competition assay using ICM-BP (Figure 3.2.2C). ICM-BP has a photoactivatable benzophenone moiety for covalent UV crosslinking with adjacent proteins, and a bioorthogonal alkyne moiety for fluorophore conjugation via click reaction.³⁷ We previously used ICM-BP for the identification of HMGBs as the target protein of **1d** and demonstrated its competitive binding toward HMGBs with **1d** in the cellular system.³⁸ With this ICM-BP in hand, we evaluated the relative binding affinity of candidate compounds **1d**, **2j**, or **2l** with HMGB1 by comparing the competitive labeling of fluorophores toward HMGB1 (Figure 3.2.2D). As shown in Figure 3.2.2E, we clearly observed the reduced labeling of Cy5 in the presence of either **1d**, **2j**, or **2l** in a dose-dependent manner. But the negative compound **2a** didn't show any dose-dependent competition, which confirms the specific competitive binding event of ICM-BP with candidate compounds. Based on this experiment, we could confirm that **2j** and **2l** bind more tightly than **1d**, which suggests that **2j** and **2l** might have enhanced activities in the regulation of HMGB1. Moreover, based on our preliminary docking simulation with HMGB1, **2j** and **2l** showed lower binding energies to HMGB1 than the original ligand **1d** did (Figure S3.2.3). Therefore, we selected **2j** and **2l** as candidate compounds with enhanced anti-inflammatory effect for further *in vivo* evaluation.

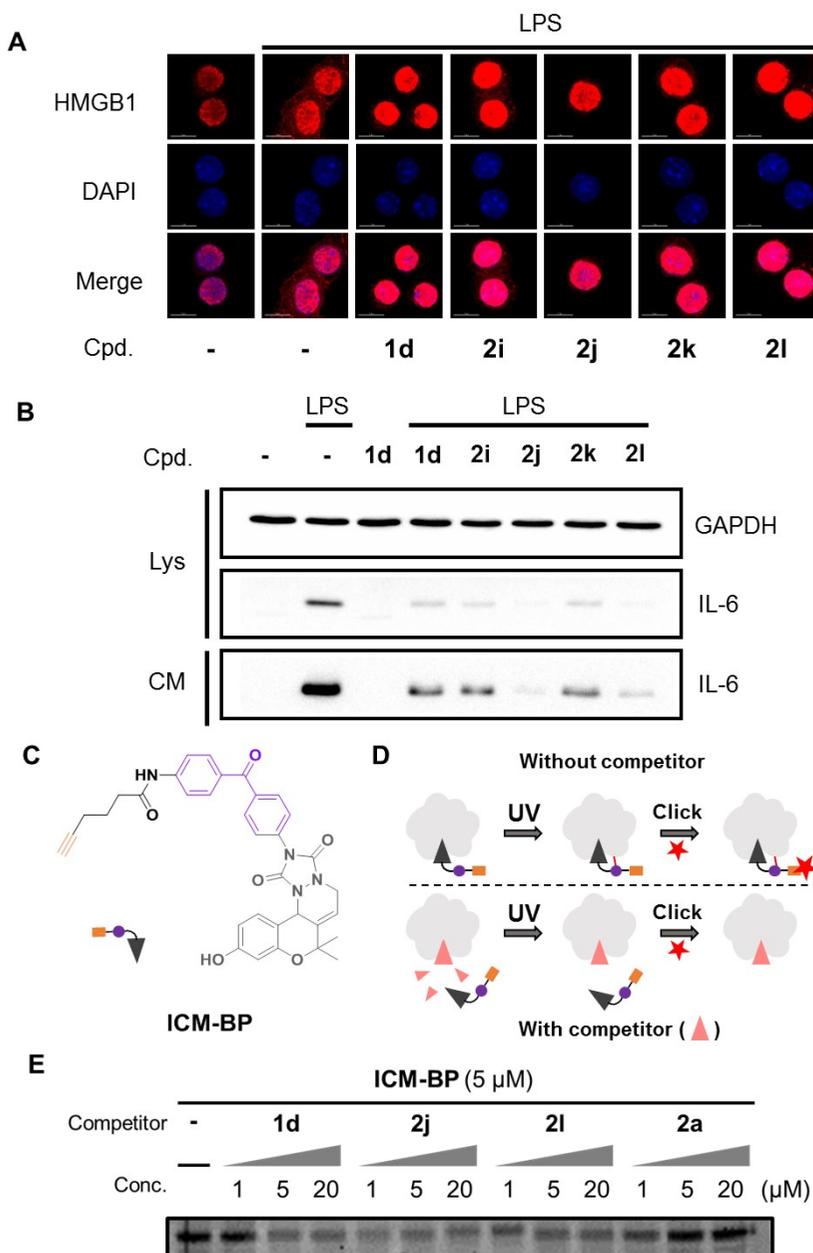


Figure 3.2.2. Inhibition of high mobility group box protein B1 (HMGB1) translocation and interleukin (IL)-6 release upon treatment with 1d, 2i, 2j, 2k, and 2l. (A) Cellular immunofluorescence images of HMGB1 after pre-treatment with 20 μ M of 1d, 2i, 2j, 2k, and 2l for 1 h, followed by 500 ng/mL lipopolysaccharide (LPS) treatment. Dimethyl sulfoxide (DMSO)

was the control. **(B)** Western blot analysis of released IL-6 after pre-treatment with 10 μ M of **1d**, **2i**, **2j**, **2k**, and **2l** for 1 h, followed by 100 ng/mL LPS treatment of Raw264.7 cells. Images shown represent at least 3 independent experiments. **(C)** Structure of ICM-BP. Purple, benzophenone moiety; Orange, alkyne moiety. **(D)** Schematic figure of *in vitro* binding competition assay. After UV crosslinking of ICM-BP, Cy5-azide was incorporated via bioorthogonal click reaction. **(E)** Fluorescence gel image of purified HMGB1 protein labeled with Cy5-azide. The same amount of HMGB1 was loaded into each lane. Competitor **1d**, **2j**, and **2l** were treated to the purified HMGB1 protein in a dose-dependent manner prior to the treatment of ICM-BP (5 μ M). **2a** was treated as a negative control, of which data exhibit no dose-dependency. Cpd., compound; Lys, cell lysate; CM, culture medium; Conc., concentration.

Inhibition of inflammatory signaling pathway by candidate compounds (1d, 2j, and 2l).

Intracellular inflammatory signaling pathways play important roles in propagating immune responses in the activated macrophage. Mitogen-activated protein kinase (MAPK) and nuclear factor kappa-light-chain enhancer in B cells (NF- κ B) complex are two of the most important inflammatory signaling systems that regulate the release of HMGBs and other cytokines in macrophages.^{4,49,50} There are three distinct types of MAPKs; p38 MAPK (p38), c-jun N-terminal kinases (JNKs), and extracellular signal-regulated kinase (ERK). Following the activation of macrophage, the MAPKs are phosphorylated and subsequently trigger the transduction of inflammatory signals within the cells.^{50,51} The western blot data showed that our compounds **1d**, **2j**, and **2l** clearly inhibited the phosphorylation of MAPK including p38, JNK, and ERK. (Figure 3.2.3A and 3.2.3B)

The NF- κ B pathway is another key signaling pathway that regulates

inflammatory responses. NF- κ B is a complex of two different protein domains, p50 and p65. In the basal state, the activity of the NF- κ B complex is inhibited by nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor (I κ B). When the activation signal reaches the macrophages, I κ B kinase (IKK) phosphorylates I κ B, which is subsequently degraded leading to the activation of the NF- κ B complex. The activated NF- κ B then translocates to the nucleus and acts as a transcription factor to produce pro-inflammatory cytokines. In this signaling pathway, the cellular translocation of the NF- κ B complex is the key event in the activation of the macrophages.^{49,52} As expected, the nuclear translocation of the NF- κ B was inhibited by treatment with our candidate compounds **1d**, **2j**, and **2l**, as evidenced by immunofluorescence assays on p65 subunit (Figure 3.2.3C and S3.2.5). Based on these observations, we concluded that our candidate compounds **1d**, **2j**, and **2l** regulated the inflammatory signaling pathways in activated macrophages.

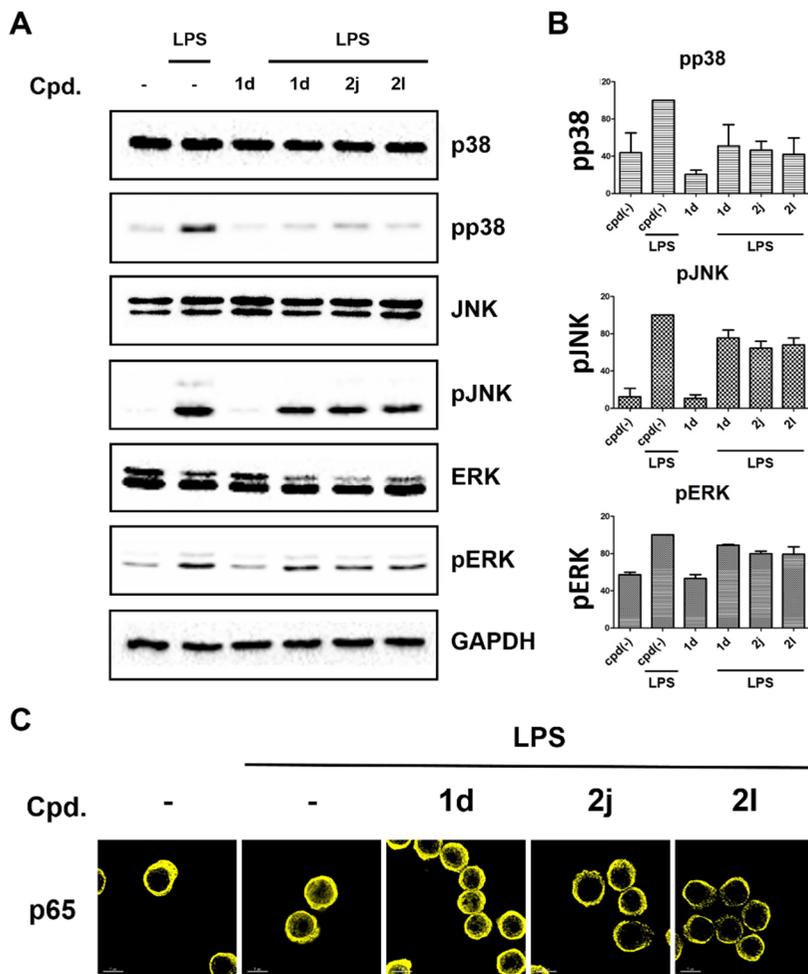


Figure 3.2.3. Inhibition of inflammatory signaling pathways in Raw264.7 cells. (A) Western blot analysis relative to the phosphorylation (p) of p38, JNK, and ERK after pre-treatment with 20 μ M **1d**, **2j**, or **2l** in Raw264.7 cells, followed by 100 ng/mL LPS treatment. DMSO was the control. (B) Quantification of phosphorylation level of p38, JNK, and ERK after compound treatment by western blot analysis. (C) Cellular NF- κ B p65 immunofluorescence images after pre-treatment with 20 μ M **1d**, **2j**, or **2l** in Raw264.7 cells, followed by 500 ng/mL LPS treatment. Data represent at least three independent experiments. Cpd., compound; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; LPS, lipopolysaccharide; DMSO, dimethyl sulfoxide; pERK, phosphorylated extracellular signal-regulated kinase; pJNK, phosphorylated c-

Jun N-terminal kinase; pp38, phosphorylated p38; NF-κB, nuclear factor kappa-light chain enhancer in B cells.

Pharmacokinetics of ICM and selected compounds.

Before investigating the selected compounds, **1d**, **2j**, and **2l** in the *in vivo* CLP-induced mouse model, we measured their liver microsomal stabilities and PK profiles (Table 3.2.3). Compared to **1d**, compounds **2j** and **2l** showed improved liver microsomal stability in both mice and humans. Probably, the PK data revealed that although the area under the curve (AUC) values of **2j** and **2l** were lower than that of **1d**, their comparable (**2j**) or longer (**2l**) half-life was expected to contribute to a better efficacy in the *in vivo* sepsis model. The **1d**, **2j**, and **2l** were all easily administered by i.p. injection and retained their activities for several hours. Along with their biological activities found in previous biological experiments, these results indicate that the compounds **1d**, **2j**, and **2l** all performed adequately in the *in vivo* mouse model following i.p. administration.

Cpd.		1d	2j	2l
Microsomal stability	Mouse (%)	0.41 ± 0.10	2.46 ± 1.83	20.2 ± 1.78
	Human (%)	0.82 ± 0.29	30.2 ± 2.81	41.2 ± 4.33
	T _{max} (h)	0.14 ± 0.10	0.25 ± 0.00	0.14 ± 0.10
PK data	C _{max} (µg/ml)	387 ± 122	100 ± 3.72	282 ± 87.0
	AUC (µg/ml*h)	327 ± 68.7	202 ± 28.1	218 ± 51.8
	T _{1/2} (h)	3.87 ± 0.84	3.45 ± 1.00	5.00 ± 0.98

Table 3.2.3. Liver microsomal stability and pharmacokinetic (PK) data of 1d, 2j, and 2l.

PK data were obtained after 10 mg/kg of each compound was administrated by intraperitoneal (i.p.) injection. Among PK parameters, C_{max}, T_{max}, AUC, and T_{1/2} were measured. Liver

microsomal stability was measured by quantifying percentage remaining compound after 30 min. Data are means \pm standard deviation (SD, n = 3). C_{\max} , maximum concentration; T_{\max} , time to achieve C_{\max} ; AUC, area under the curve; $T_{1/2}$, half-life.

Enhanced *in vivo* therapeutic effect of **2j and **2l** in CLP-induced mouse model of sepsis.**

Finally, we compared the *in vivo* therapeutic effect of **1d**, **2j**, and **2l** in a CLP-induced mouse model. The surgical procedure used to perform the CLP is illustrated in Figure 3.2.4A. Briefly, after excising the abdomen, the exposed cecum was ligated just below the ileocecal valve and then punctured using a surgical needle. Then, the excised abdomen was closed. The CLP-induced mice were treated with our candidate compounds for 21 days by daily i.p. injection observing that the survival rate of the **1d**-, **2j**-, and **2l**-treated mice was significantly improved compared to that of the vehicle-treated mice (Figure 3.2.4B). Both compounds **2j** and **2l** showed better efficacy than **1d** in improving the survival rate, being **2j** the most active molecule. Compared to **1d** with a 45% survival rate, 74% and 56% were found for **2j**- and **2l**-treated mice, respectively, while only 10% of the vehicle-treated mice survived. Furthermore, serum IL-6 levels were significantly decreased in the survived mice (Figure S3.2.6), which is consistent with our previous observation in Figure 3.2.2C and Figure S3.2.1. Therefore, we successfully enhanced the therapeutic effect of **1d** by structural optimization obtaining **2j** and **2l** as potential drug candidates for the treatment of sepsis.

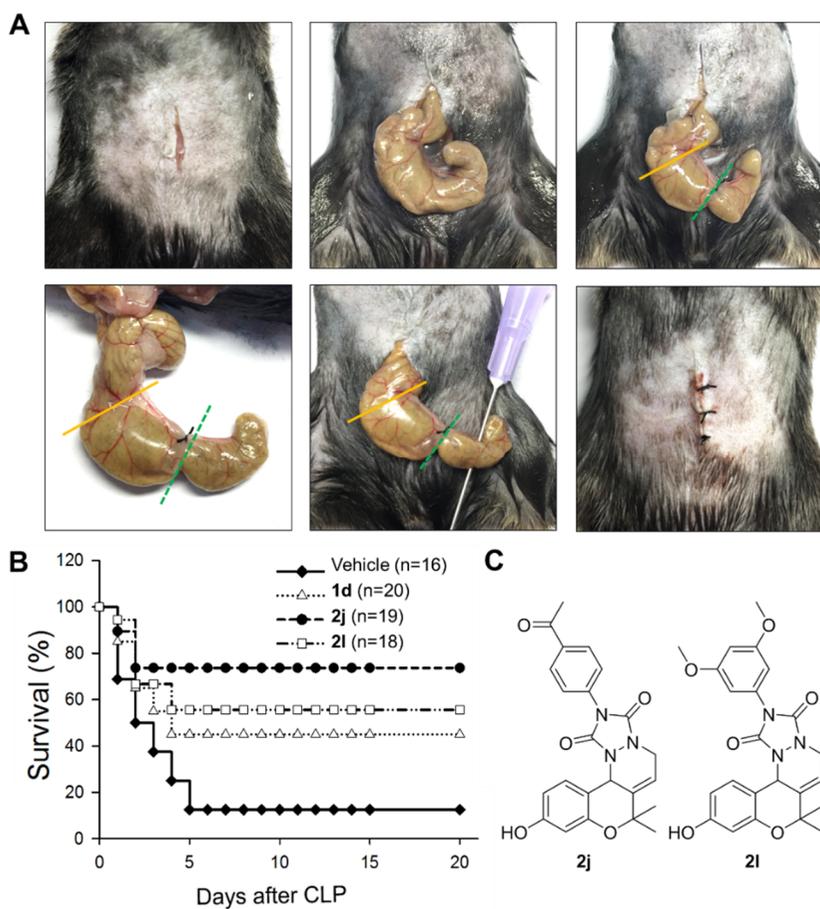


Figure 3.2.4. Therapeutic effect of selected compounds in in vivo CLP-induced mouse model. (A) Cecal Ligation and puncture procedure. Severity of sepsis can be controlled by the position of ligation. (B) Percentage survival rate of vehicle- and compound **1d**-, **2j**-, and **2l**-treated CLP-induced mice. **1d**, **2j**, and **2l** were administered intraperitoneally in a daily dose of 10 mg/kg for 21 days, starting from 1 h prior to CLP induction. Data shown is the accumulation of 3 independent experiments, each trial with at least N = 5. (C) Chemical structure of **2j** and **2l**.

3.2.4. Conclusion

In this study, we explored the therapeutic efficacy in the treatment of sepsis of a previously reported small molecule, ICM (**1d**), which regulated the secretion of HMGB1 and HMGB2 in activated microglia.³⁸ Pro-inflammatory cytokines, such as TNF- α and IL-1 β , have shown limited clinical usefulness because of their characteristic early-stage mediator effects,^{7,19-21} while the clinical relevance of antagonizing HMGBs in sepsis has been proposed and reported in several studies.³⁰⁻³⁴ Indeed, as late-stage mediators of systemic inflammatory responses, HMGBs regulate the production of other pro-inflammatory cytokines, thereby modulating the systemic inflammatory response. Therefore, the inhibition of HMGB1 release could downregulate the production of pro-inflammatory cytokines.^{30,31} With this in mind, we evaluated our novel HMGB1-regulating compound, ICM (**1d**) in an *in vivo* CLP-induced mouse model finding that it efficiently ameliorated the sepsis pathogenesis.

We further sought to improve the efficacy of **1d** by synthesizing 30 analogues and their anti-inflammatory activities were evaluated by monitoring the inhibition of NO release in the Griess assay. Based on our structure-activity relationship study and secondary biological evaluations, we selected **2j** and **2l**, with different substituents on the **1d** phenyl ring, as our final candidates. Both **2j** and **2l** inhibited the translocation of HMGB1 from the nucleus to the cytosol, improved the inhibition of IL-6 release, and regulated the downstream inflammatory signaling pathways. Furthermore, **2j** and **2l** also showed better liver microsomal stabilities and reasonable PK profiles compared to that of **1d**. Finally, the daily *i.p.* administration of **2j** and **2l** in the *in vivo* CLP-induced

mouse model showed enhanced therapeutic effects, and **2j** exhibited the best therapeutic efficacy against sepsis. Therefore, this research study has provided considerable evidence supporting the use of selective regulation of HMGB1 release as a promising therapeutic strategy for the treatment of sepsis by using small molecules.

3.2.5. Experimental Section

General Information of Biological Experiments

1. Kits, Reagents and Materials.

Dulbecco's modified eagle medium (DMEM), fetal bovine serum (FBS), and antibiotic-antimycotic solution were purchased from Gibco, Invitrogen. Phosphate-buffered saline (PBS) buffer were purchased from WELGENE. Raw264.7 cell line was obtained from Korean Cell Line Bank. Compounds were prepared in dimethyl sulfoxide (DMSO) solution for biological evaluation. DMSO was purchased from Acros Organics. Lipopolysaccharide (LPS) from *Escherichia Coli* 055:B5 purified by phenol extraction, *N*-(1-naphthyl)ethylenediamine dihydrochloride, sulfanilamide, phosphoric acid (85%), paraformaldehyde, tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]-amine (TBTA), CuSO₄ and *t*BuOH were purchased from Sigma-Aldrich. Tris(2-carboxyethyl)-phosphine (TCEP) was purchased from TCI. Cy5-azide was purchased from Lumiprobe. Hoechst 33342 was purchased from Life Technologies. Ez-cytox WST assay reagent was purchased from Daeil-Bio. To measure the protein concentration of cell lysate, Micro BCA™ protein assay kit was purchased from PIERCE. Mouse TNF- α DuoSet ELISA kit was purchased from R&D Biosystems. 3,3',5,5"-tetramethylbenzidine (TMB), a substrate of horseradish peroxidase (HRP) for TNF- α ELISA, was purchased from Invitrogen. For the quantification of serum IL-6 level by ELISA assay, streptavidin-conjugated horseradish peroxidase and TMB solution was

purchased from eBioscience. Protein gel casters and western blot equipment including polyvinylidene difluoride (PVDF) membrane were purchased from Bio-Rad. Amersham ECL prime western blotting detection system (Amersham ECL prime solution) was purchased from GE Healthcare Life Science. 100 mm cell culture dish, transparent 96-well plate [#3598] and transparent 6-well plate [#3506] were purchased from CORNING. Nunc™ Lab-Tek™ II chambered coverglass was purchased from Thermo Scientific. ELISA plates were purchased from Costar.

2. Antibodies and protein

Anti-high mobility group box protein 1 (HMGB1) rabbit polyclonal antibody, Anti-interleukin 6 (IL-6) rabbit polyclonal antibody, and goat anti-rabbit TRITC (IgG H&L) antibody were purchased from Abcam. Anti-p44/42 MAPK (ERK) rabbit monoclonal antibody, anti-phospho-p44/42 MAPK (Thr202/Tyr204) (pERK) rabbit polyclonal antibody, anti-SAPK/JNK (JNK) rabbit polyclonal antibody, anti-phospho-SAPK/JNK (Thr183/Tyr185) (pJNK) rabbit polyclonal antibody, anti-p38 MAPK (p38) rabbit polyclonal antibody, anti-phospho-p38 MAPK (Thr180/Tyr182) (pp38) rabbit polyclonal antibody, anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) rabbit monoclonal antibody, and HRP-labeled anti-rabbit secondary antibody were purchased from Cell Signaling Technology. Biotinylated anti-IL-6 monoclonal antibody for ELISA assay was purchased from eBioscience. Purified recombinant human HMGB1 protein was purchased from R&D Biosystems.

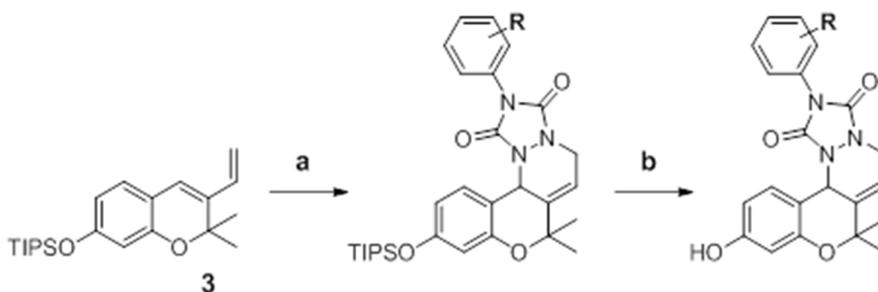
3. Instruments and Programs

Measurement of absorbance in 96-well plate for Griess assay, WST assay, TNF- α ELISA assay, and BCA assay was performed with BioTek Synergy HT Microplate reader. Measurement of absorbance for serum IL-6 ELISA assay was performed with a Multiskan Ascent (Labsystems, Kennett Square, PA). For the analysis of chemiluminescent signal for western blotting assay, ChemiDoc™ MP imaging system from Bio-Rad was used. Quantification of signal was done by ImageLab 4.0 program provided by Bio-Rad. The in-gel fluorescence labeling signal was detected and visualized with Typhoon Trio from Amersham Bioscience. Cy5 channel (633 nm excitation) was used for the visualization. Data was analyzed by ImageQuant TL program from Amersham Bioscience. BLAK-Ray (B-100AP) UV lamp was purchased from UVP for the irradiation of 365 nm UV light. DeltaVision Elite imaging system from GE Healthcare was used for immunofluorescence imaging. Objective lenses were equipped with Olympus IX-71 inverted microscope with PLAN APO 60x/Oil (PLAPON60xO), 1.42 NA, WD 0.15 mm. sCMOS camera and InSightSSI fluorescence illumination module were equipped with the system. Four-color standard filter set [GE Healthcare, 52-852113-003] was used for imaging. For live cell imaging, CO₂ supporting chamber with an objective air heater were installed with the system. Images were analyzed with SoftWorks program supported by GE Healthcare. Provided graphs were analyzed with GraphPad Prism 5 software.

Synthetic Scheme and Characterization of Compound 2e, 2j, 2k, 2l, 2m and 2o

2o

■ Synthetic scheme for the preparation of 1d, 2a–2o from 3

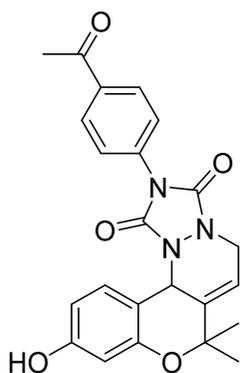


Reaction conditions: a) 4-Substituted-1,2,4-triazoline-3,5-dione, toluene/THF(1:1), r. t., 1 h to 6 h; b) HF/pyridine/THF(5:5:90), r. t., 6 h.

■ 10-Hydroxy-2-(2-methoxyphenyl)-7,7-dimethyl-5,12b-dihydro-1H,7H-chromeno[4,3-c][1,2,4]triazolo[1,2-a]pyridazine-1,3(2H)-dione (2e, 2-methoxy-ICM)

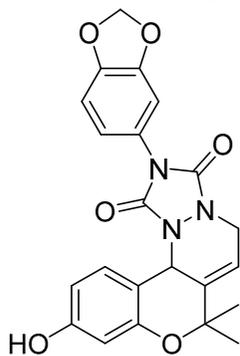
$R_f = 0.25$ (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); $^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 9.47 (s, 1H), 7.53 (t, $J = 7.8$ Hz, 1H), 7.41 (d, $J = 8.0$ Hz, 1H), 7.26 (d, $J = 8.8$ Hz, 1H), 7.11 (t, $J = 7.6$ Hz, 1H), 6.91 (d, $J = 8.4$ Hz, 1H), 6.38–6.42 (m, 1H), 6.27 (s, 1H), 5.93 (s, 1H), 5.58 (s, 1H), 4.02–4.24 (m, 2H), 3.88 (s, 3H), 1.55 (s, 3H), 1.51 (s, 3H) (more stable conformer); $^{13}\text{C NMR}$ (100 MHz, DMSO- d_6) δ 158.2, 158.2, 155.4, 154.0, 153.9, 151.4, 136.5, 131.2, 130.4, 123.7, 120.7, 119.4, 114.8, 112.5, 104.3, 104.2, 79.2, 56.0, 49.3, 44.1, 27.3, 26.8; LRMS(ESI $^+$) m/z calcd for $\text{C}_{22}\text{H}_{22}\text{N}_3\text{O}_5$ [M+H] $^+$ 408.16; Found 408.2.

- 2-(4-Acetylphenyl)-10-hydroxy-7,7-dimethyl-5,12b-dihydro-1*H*,7*H*-chromeno[4,3-*c*][1,2,4]triazolo[1,2-*a*]pyridazine-1,3(2*H*)-dione (**2j**, 4-acetyl-*ICM*)



R_f = 0.26 (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 9.47 (s, 1H), 8.13 (d, J = 8.0 Hz, 2H), 7.79 (d, J = 8.8 Hz, 2H), 6.87 (d, J = 8.4 Hz, 1H), 6.37 (dd, J = 8.8 Hz, 2.0 Hz, 1H), 6.25 (d, J = 1.6 Hz, 1H), 5.95 (s, 1H), 5.61 (s, 1H), 4.11–4.25 (m, 2H), 2.64 (s, 3H), 1.54 (s, 3H), 1.51 (s, 3H); ^{13}C NMR (100 MHz, $\text{DMSO-}d_6$) δ 197.2, 158.2, 153.8, 153.3, 150.5, 136.0, 135.8, 135.6, 128.9, 125.7, 124.5, 114.6, 113.6, 108.6, 104.1, 79.0, 49.3, 44.0, 27.4, 26.9, 26.8; LRMS(ESI⁺) m/z calcd for $\text{C}_{23}\text{H}_{22}\text{N}_3\text{O}_5$ [M+H]⁺ 420.16; Found 420.1. In vitro solubility (μSOL) = $435.1 \pm 5.06 \mu\text{M}$.¹

- 2-(Benzo[*d*][1,3]dioxol-5-yl)-10-hydroxy-7,7-dimethyl-5,12b-dihydro-1*H*,7*H*-chromeno[4,3-*c*][1,2,4]triazolo[1,2-*a*]pyridazine-1,3(2*H*)-dione (**2k**, 3,4-(methylenedioxy)-*ICM*)

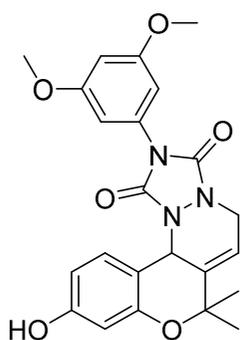


3,4-(methylenedioxy)-*ICM*)

R_f = 0.07 (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); ^1H NMR (400 MHz, CDCl_3) δ 7.00–7.04 (m, 2H), 6.83–6.92 (m, 2H), 6.31–6.36 (m, 2H), 6.07 (s, 1H), 6.02 (s, 2H), 5.78–5.81 (m, 1H), 5.66 (s, 1H), 4.24–4.30 (m, 1H), 4.04–4.10 (m, 1H), 1.78 (s,

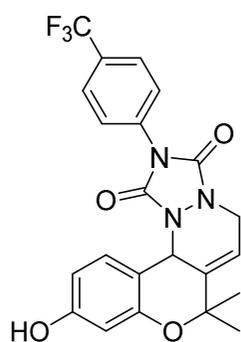
6H); ^{13}C NMR (100 MHz, CDCl_3) δ 157.2, 154.3, 154.1, 151.8, 148.3, 147.9, 137.4, 124.6, 124.4, 120.1, 115.4, 112.7, 109.0, 108.6, 107.4, 105.0, 102.0, 78.9, 50.2, 44.3, 27.6, 26.9; LRMS(ESI^+) m/z calcd for $\text{C}_{22}\text{H}_{20}\text{N}_3\text{O}_6$ $[\text{M}+\text{H}]^+$ 422.14; Found 422.1.

- 2-(3,5-Dimethoxyphenyl)-10-hydroxy-7,7-dimethyl-5,12b-dihydro-1*H*,7*H*-chromeno[4,3-*c*][1,2,4]triazolo[1,2-*a*]pyridazine-1,3(2*H*)-dione (**2l**, 3,5-dimethoxy-*ICM*)



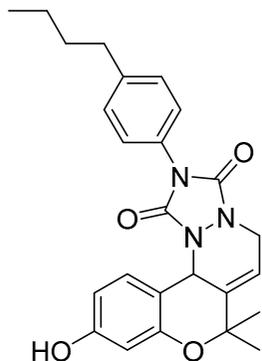
R_f = 0.08 (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 9.46 (s, 1H), 6.84 (d, J = 8.4 Hz, 1H), 6.77 (d, J = 2.4 Hz, 2H), 6.614 (s, 1H), 6.38 (dd, J = 8.0 Hz, 2.2 Hz, 1H), 6.25 (d, J = 2.4 Hz, 1H), 5.94 (s, 1H), 5.57 (s, 1H), 4.07–4.22 (m, 2H), 3.79 (s, 6H), 1.53 (s, 3H), 1.50 (s, 3H); ^{13}C NMR (100 MHz, $\text{DMSO-}d_6$) δ 160.4, 158.2, 153.8, 153.7, 150.8, 136.1, 133.0, 124.4, 114.6, 113.6, 108.6, 105.0, 79.0, 55.6, 55.5, 49.2, 44.0, 27.3, 26.7; LRMS(ESI^+) m/z calcd for $\text{C}_{23}\text{H}_{24}\text{N}_3\text{O}_6$ $[\text{M}+\text{H}]^+$ 438.17; Found 438.1. In vitro solubility (μSOL) = 166.3 ± 6.26 μM .¹

- 10-Hydroxy-7,7-dimethyl-2-(4-(trifluoromethyl)phenyl)-5,12b-dihydro-1*H*,7*H*-chromeno[4,3-*c*][1,2,4]triazolo[1,2-*a*]pyridazine-1,3(2*H*)-dione (**2m**, 4-(trifluoromethyl)-*ICM*)



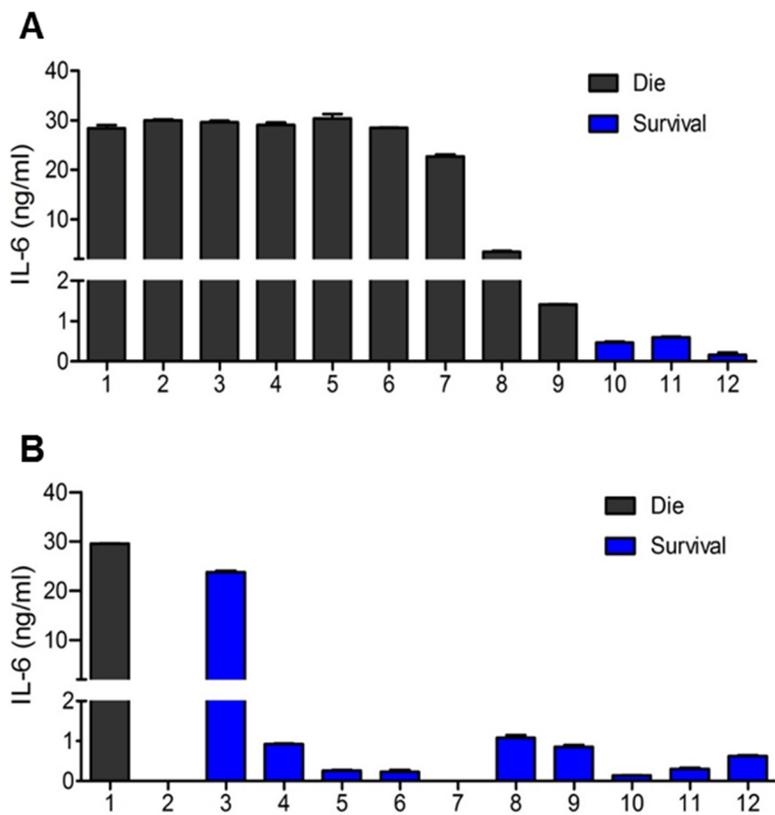
R_f = 0.21 (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); ^1H NMR (400 MHz, $\text{DMSO-}d_6$) δ 9.47 (brs, 1H), 7.95 (d, J = 8.4 Hz, 2H), 7.89 (d, J = 8.8 Hz, 2H), 6.88 (d, J = 8.8 Hz, 1H), 6.37 (dd, J = 8.4 Hz, 1.6 Hz, 1H), 6.25 (d, J = 2.0 Hz, 1H), 5.95 (s, 1H), 5.61 (s, 1H), 4.16–4.23 (m, 2H), 1.54 (s, 3H), 1.51 (s, 3H); ^{13}C NMR (100 MHz, $\text{DMSO-}d_6$) δ 158.2, 153.8, 153.2, 150.4, 130.9, 135.2, 126.4, 126.1, 124.5, 114.5, 113.5, 108.6, 104.1, 79.2, 79.0, 49.3, 44.0, 27.4, 26.8; LRMS(ESI⁺) m/z calcd for $\text{C}_{22}\text{H}_{19}\text{F}_3\text{N}_3\text{O}_4$ $[\text{M}+\text{H}]^+$ 446.13; Found 446.0.

■ 2-(4-Butylphenyl)-10-hydroxy-7,7-dimethyl-5,12b-dihydro-1H,7H-chromeno[4,3-c][1,2,4]triazolo[1,2-a]pyridazine-1,3(2H)-dione (**2o**, 4-*n*-butyl-ICM)

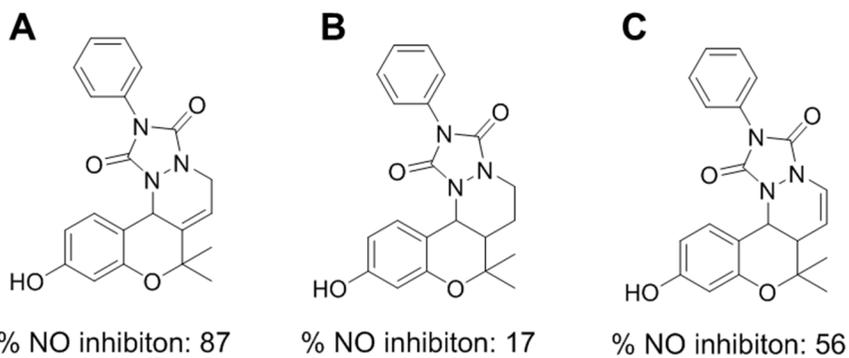


R_f = 0.63 (1:1 = EtOAc:n-hexane, v/v, with anisaldehyde staining); ^1H NMR (400 MHz, CDCl_3) δ 7.48 (d, J = 8.0 Hz, 2H), 7.31 (d, J = 8.4 Hz, 2H), 6.88 (d, J = 8.0 Hz, 1H), 6.32–6.37 (m, 2H), 5.78–5.81 (m, 2H), 5.68 (s, 1H), 4.28 (dd, J = 16.4 Hz, 4.4 Hz, 1H), 4.05–4.10 (m, 1H), 2.65 (t, J = 7.6 Hz, 2H), 1.55–1.64 (m, 2H), 1.56 (s, 6H), 1.37 (sextet, J = 7.5 Hz, 2H), 0.94 (t, J = 7.2 Hz, 2H); ^{13}C NMR (100 MHz, CDCl_3) δ 157.1, 154.3, 154.1, 151.8, 143.6, 137.5, 129.4, 128.6, 125.5, 124.7, 115.6, 112.8, 109.0, 105.0, 79.0, 50.2, 44.3, 35.5, 33.6, 27.6, 26.9, 22.4, 14.1; LRMS(ESI⁺) m/z calcd for $\text{C}_{25}\text{H}_{28}\text{N}_3\text{O}_4$ $[\text{M}+\text{H}]^+$ 434.21; Found 434.2.

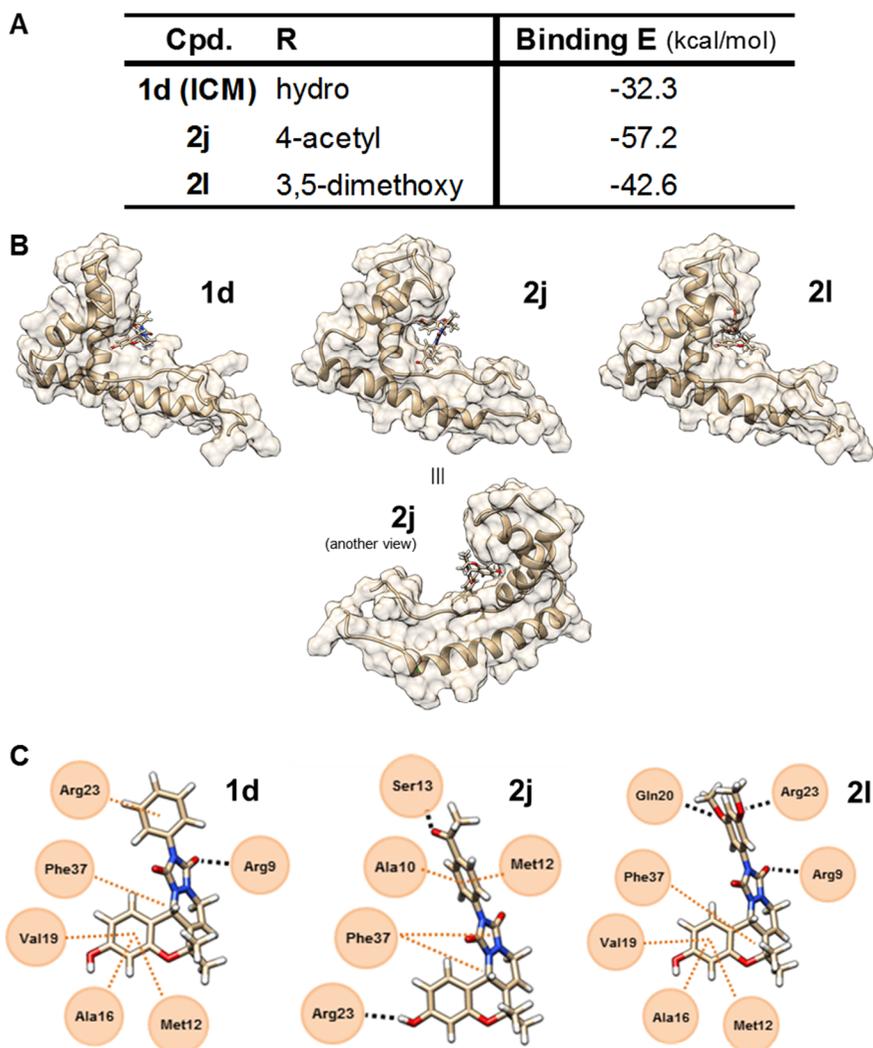
3.2.6. Supporting Information



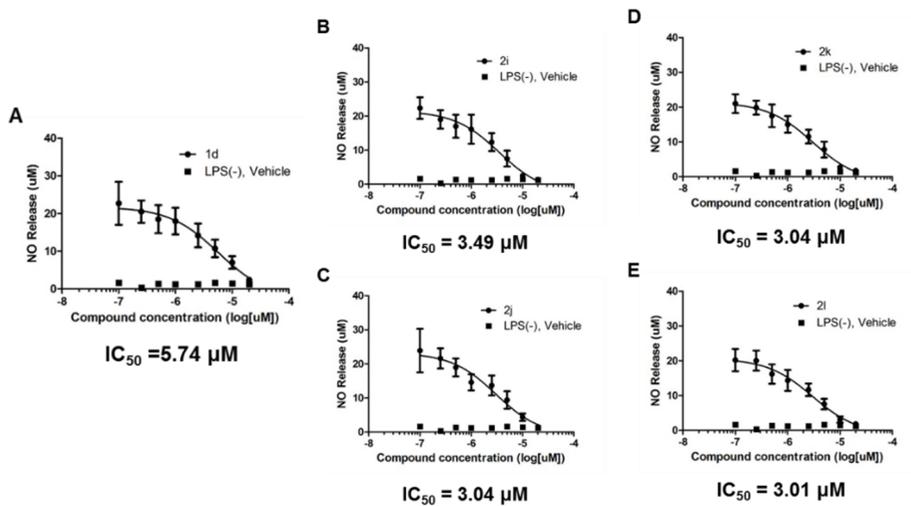
Supplementary Figure S3.2.1. Serum IL-6 level of *in vivo* CLP-induced mouse model. (A) Serum IL-6 level in vehicle-treated and (B) **1d**-treated mice. IL-6 levels in survived mice were decreased compared to died mice. Compound **1d** were treated in a daily dose of 10 mg/kg for 9 days, starting from 1 h before CLP induction.



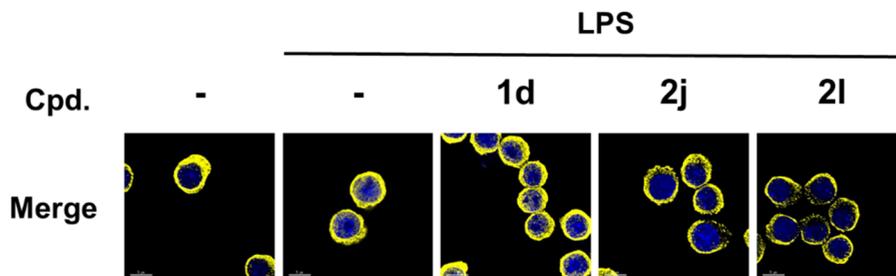
Supplementary Figure S3.2.2. The removal or shift of the double bond of **1d** reduces the inhibitory effect towards nitric oxide release. (A) Chemical structure of **1d**. Double bond removal (B) or double bond migration (C) reduces the % NO release inhibition to 17% and 56%.



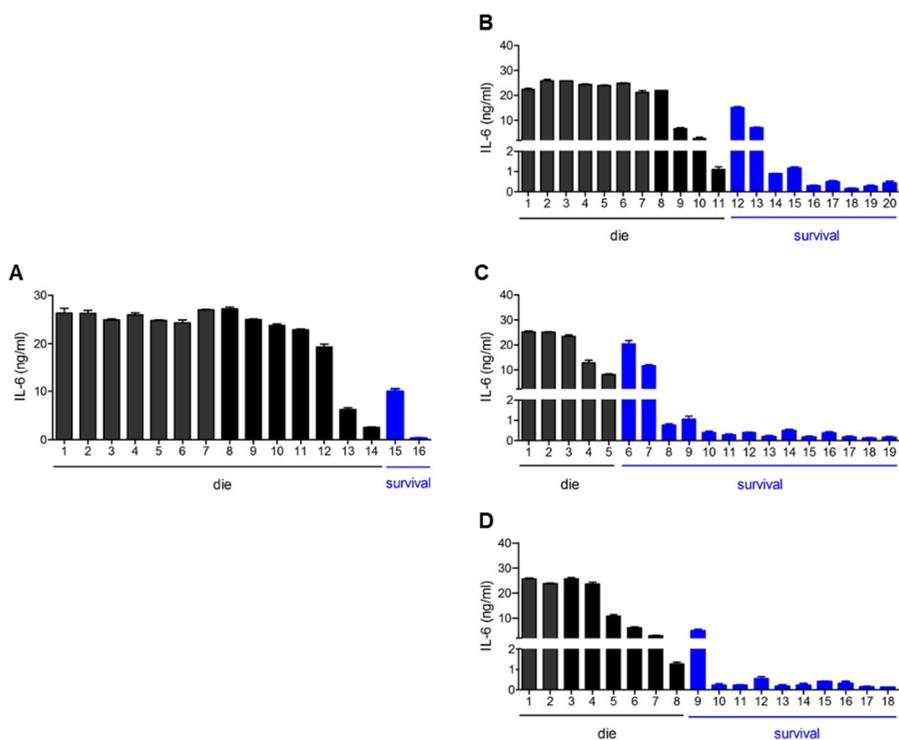
Supplementary Figure S3.2.3. Protein-ligand docking simulation of compounds **1d**, **2j** and **2l** with HMGB1 protein (PDB ID: 1HMF); (A) Binding energies of **1d**, **2j** and **2l** as ligands of HMGB1. **2j** and **2l** bind to HMGB1 better than **1d**; (B) Binding modes of **1d**, **2j** and **2l**. **2j** showed higher docking score when it binds to HMGB1 in the same binding site with a different binding mode; (C) Interaction of **1d**, **2j** and **2l** with adjacent amino acids at the binding cavity of HMGB1. Amino acid residues within the range of 3 Å from ligands were labeled around each ligand. Orange dotted lines are the nearest hydrophobic interactions and black dotted lines are hydrogen bondings.



Supplementary Figure S3.2.4. Dose-dependent inhibition of nitric oxide release by compounds **1d** (A), **2i** (B), **2j** (C), **2k** (D) and **2l** (E). Raw264.7 macrophage cells were treated with 100 ng/mL LPS, followed by the treatment with each compound **1d**, **2i**, **2j**, **2k** and **2l** in various concentrations indicated in graph.



Supplementary Figure S3.2.5. Merged TRITC, DAPI channel images of p65 immunofluorescence. DAPI channel image indicates nucleus, stained by Hoechst 33342 dye.



Supplementary Figure S3.2.6. Serum IL-6 levels of *in vivo* CLP-induced mouse model. Serum IL-6 level in vehicle-treated (A), **1d**- (B), **2j**- (C), and **2l**-treated mice (D). IL-6 levels in survived mice were decreased compared to dead mice. Compounds **1d**, **2j**, and **2l** were treated in a daily dose of 10 mg/kg for 21 days, starting from 1 h before CLP induction.

3.2.7. References

- (1) Bone, R. C.; Sibbald, W. J.; Sprung, C. L. *Chest* **1992**, *101*, 1481–1483.
- (2) Levy, M. M.; Fink, M. P.; Marshall, J. C.; Abraham, E.; Angus, D.; Cook, D.; Cohen, J.; Opal, S. M.; Vincent, J. L.; Ramsay, G. *Crit. Care Med.* **2003**, *31*, 1250–1256.

- (3) Deutschman, C. S.; Tracey, K. J. *Immunity* **2014**, *40*, 463–475.
- (4) Takeuchi, O.; Akira, S. *Cell* **2010**, *140*, 805–820.
- (5) Mogensen, T. H. *Clin. Microbiol. Rev.* **2009**, *22*, 240–273.
- (6) Rittirsch, D.; Flierl, M. A.; Ward, P. A. *Nat. Rev. Immunol.* **2008**, *8*, 776–787.
- (7) Angus, D. C.; van der Poll, T. *N. Engl. J. Med.* **2013**, *369*, 840–851.
- (8) Angus, D. C.; Linde-Zwirble, W. T.; Lidicker, J.; Clermont, G.; Carcillo, J.; Pinsky, M. R. *Crit. Care Med.* **2001**, *29*, 1303–1310.
- (9) Rangel-Frausto, M. S.; Pittet, D.; Costigan, M.; Hwang, T.; Davis, C. S.; Wenzel, R. P. *JAMA* **1995**, *273*, 117–123.
- (10) Kaukonen, K.-M.; Bailey, M.; Suzuki, S.; Pilcher, D.; Bellomo, R. *JAMA* **2014**, *311*, 1308–1316.
- (11) Torio, C. M.; Andrews, R. M. *National Inpatient Hospital Costs: The Most Expensive Conditions by Payer, 2011*; Agency for Health Care Policy and Research (US): Rockville (MD), **2013**.
- (12) Fink, M. P.; Warren, H. S. *Nat. Rev. Drug Discovery* **2014**, *13*, 741–758.
- (13) Ulloa, L.; Tracey, K. J. *Trends Mol. Med.* **2005**, *11*, 56–63.
- (14) Ertel, W.; Kremer, J. P.; Kenney, J.; Steckholzer, U.; Jarrar, D.; Trentz, O.; Schildberg, F. W. *Blood* **1995**, *85*, 1341–1347.
- (15) Whang, K. T.; Vath, S. D.; Becker, K. L.; Snider, R. H.; Nylen, E. S.; Muller, B.; Li, Q.; Tamarkin, L.; White, J. C. *Shock* **2000**, *14*, 73–78.
- (16) Bouchon, A.; Facchetti, F.; Weigand, M. A.; Colonna, M. *Nature* **2001**, *410*, 1103–1107.

- (17) Lorente, J. A.; Marshall, J. C.. *Shock* **2005**, *24*, 107–119.
- (18) Alexander, H. R.; Doherty, G. M.; Buresh, C. M.; Venzon, D. J.; Norton, J. A. *J. Exp. Med.* **1991**, *173*, 1029–1032.
- (19) Abraham, E.; Laterre, P.-F.; Garbino, J.; Pingleton, S.; Butler, T.; Dugernier, T.; Margolis, B.; Kudsk, K.; Zimmerli, W.; Anderson, P.; Reynaert, M.; Lew, D.; Lesslauer, W.; Passe, S.; Cooper, P.; Burdeska, A.; Modi, M.; Leighton, A.; Salgo, M.; Van der Auwera, P. *Crit. Care Med.* **2001**, *29*, 503–510.
- (20) Opal, S. M.; Fisher, C. J. J.; Dhainaut, J.-F. A.; Vincent, J.-L.; Brase, R.; Lowry, S. F.; Sadoff, J. C.; Slotman, G. J.; Levy, H.; Balk, R. A.; Shelly, M. P.; Pribble, J. P.; LaBrecque, J. F.; Lookabaugh, J.; Donovan, H.; Dubin, H.; Baughman, R.; Norman, J.; DeMaria, E.; Matzel, K.; Abraham, E.; Seneff, M. *Crit. Care Med.* **1997**, *25*, 1115–1124.
- (21) Kellum, J. A.; Kong, L.; Fink, M. P.; Weissfeld, L. A.; Yealy, D. M.; Pinsky, M. R.; Fine, J.; Krichevsky, A.; Delude, R. L.; Angus, D. C. *Arch. Intern. Med.* **2007**, *167*, 1655–1663.
- (22) Wang, H.; Bloom, O.; Zhang, M.; Vishnubhakat, J. M.; Ombrellino, M.; Che, J.; Frazier, A.; Yang, H.; Ivanova, S.; Borovikova, L.; Manogue, K. R.; Faist, E.; Abraham, E.; Andersson, J.; Andersson, U.; Molina, P. E.; Abumrad, N. N.; Sama, A.; Tracey, K. J. *Science* **1999**, *285*, 248–251.
- (23) Lotze, M. T.; Tracey, K. J. *Nat. Rev. Immunol.* **2005**, *5*, 331–342.
- (24) Yang, H.; Tracey, K. J. *Biochim. Biophys. Acta* **2010**, *1799*, 149–156.

- (25) Erlandsson Harris, H.; Andersson, U. *Eur. J. Immunol.* **2004**, *34*, 1503–1512.
- (26) Sims, G. P.; Rowe, D. C.; Rietdijk, S. T.; Herbst, R.; Coyle, A. J. *Annu. Rev. Immunol.* **2010**, *28*, 367–388.
- (27) Liu, X.-X.; Wang, C.; Huang, S.-F.; Chen, Q.; Hu, Y.-F.; Zhou, L.; Gu, Y. *Sci. Rep.* **2016**, *6*, 24073.
- (28) Rittirsch, D.; Huber-Lang, M. S.; Flierl, M. A.; Ward, P. A. *Nat. Protoc.* **2009**, *4*, 31–36.
- (29) Hubbard, W. J.; Choudhry, M.; Schwacha, M. G.; Kerby, J. D.; Rue, L. W.; Bland, K. I.; Chaudry, I. H. *Shock* **2005**, *24*, 52–57.
- (30) Yang, H.; Ochani, M.; Li, J.; Qiang, X.; Tanovic, M.; Harris, H. E.; Susarla, S. M.; Ulloa, L.; Wang, H.; DiRaimo, R.; Czura, C. J.; Wang, H.; Roth, J.; Warren, H. S.; Fink, M. P.; Fenton, M. J.; Andersson, U.; Tracey, K. J. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 296–301.
- (31) Wang, H.; Liao, H.; Ochani, M.; Justiniani, M.; Lin, X.; Yang, L.; Al-Abed, Y.; Wang, H.; Metz, C.; Miller, E. J.; Tracey, K. J.; Ulloa, L. *Nat. Med.* **2004**, *10*, 1216–1221.
- (32) Wang, H.; Ward, M. F.; Sama, A. E. *Expert Opin. Ther. Targets* **2014**, *18*, 257–268.
- (33) Li, W.; Ashok, M.; Li, J.; Yang, H.; Sama, A. E.; Wang, H. *PLoS One* **2007**, *2*, e1153.
- (34) Li, W.; Zhu, S.; Li, J.; Assa, A.; Jundoria, A.; Xu, J.; Fan, S.; Eissa, N. T.; Tracey, K. J.; Sama, A. E.; Wang, H. *Biochem. Pharmacol.* **2011**, *81*, 1152–1163.

- (35) Kim, J.; Kim, H.; Park, S. B. *J. Am. Chem. Soc.* **2014**, *136*, 14629–14638.
- (36) Zhu, M.; Lim, B. J.; Koh, M.; Park, S. B. *ACS Comb. Sci.* **2012**, *14*, 124–134.
- (37) Park, J.; Oh, S.; Park, S. B. *Angew. Chem. Int. Ed.* **2012**, *51*, 5447–5451.
- (38) Lee, S.; Nam, Y.; Koo, J. Y.; Lim, D.; Park, J.; Ock, J.; Kim, J.; Suk, K.; Park, S. B. *Nat. Chem. Biol.* **2014**, *10*, 1055–1060.
- (39) Mosser, D. M.; Edwards, J. P. *Nat. Rev. Immunol.* **2008**, *8*, 958–969.
- (40) Filiano, A. J.; Gadani, S. P.; Kipnis, J. *Brain Res.* **2015**, *1617*, 18–27.
- (41) Bonaldi, T.; Talamo, F.; Scaffidi, P.; Bachi, A.; Rubartelli, A.; Agresti, A.; Bianchi, M. E. *EMBO J.* **2003**, *22*, 5551–5560.
- (42) Ghosh, M.; Garcia-Castillo, D.; Aguirre, V.; Golshani, R.; Atkins, C. M.; Bramlett, H. M.; Dietrich, W. D.; Pearse, D. D. *Glia* **2012**, *60*, 1839–1859.
- (43) Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B.-S.; Lee, H.; Lee, J.-O. *Nature* **2009**, *458*, 1191–1195.
- (44) Qiao, S.; Luo, Q.; Zhao, Y.; Zhang, X. C.; Huang, Y. *Nature* **2014**, *511*, 108–111.
- (45) Dong, H.; Xiang, Q.; Gu, Y.; Wang, Z.; Paterson, N. G.; Stansfeld, P. J.; He, C.; Zhang, Y.; Wang, W.; Dong, C. *Nature* **2014**, *511*, 52–56.
- (46) Fanying, M.; Lowell, C. A. *J. Exp. Med.* **1997**, *185*, 1661–1670.
- (47) Hajishengallis, G.; Lambris, J. D. *Nat. Rev. Immunol.* **2011**, *11*, 187–200.

- (48) Tsikas, D. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2007**, *851*, 51–70.
- (49) Kawai, T.; Akira, S. *Trends Mol. Med.* **2007**, *13*, 460–469.
- (50) Zhang, W.; Liu, H. T. *Cell Res.* **2002**, *12*, 9–18.
- (51) Rao, K. M. K. *J. Leukoc. Biol.* **2001**, *69*, 3–10.
- (52) Oeckinghaus, A.; Hayden, M. S.; Ghosh, S. *Nat. Immunol.* **2011**, *12*, 695–708.
- (53) Chang, C. E.; Gilson, M. K. *J. Comput. Chem.* **2003**, *24*, 1987–1998.
- (54) Zhu, M.; Kim, M. H.; Lee, S.; Bae, S. J.; Kim, S. H.; Park, S. B. *J. Med. Chem.* **2010**, *53*, 8760–8764.

국문 초록

생명 현상의 탐구를 위한 새로운 바이오프로브 및 생리활성 저분자 화합물의 합리적 디자인 및 최적화

인간 생명체의 크고 작은 변화들에 있어서, 모든 기초적인 영역 내면에 화학적 상호작용이 그 기저를 이루고 있다. 눈으로 볼 수 없는 몸 속의 작은 세계를 관찰하는데 있어서, 생체 시스템을 탐지하거나 생체 시스템에 변화를 가져올 수 있는 화학적 도구들의 개발은 필수불가결하다. 다양한 질병 치료제들을 포함해, 저분자 화합물은 인간을 포함한 유기체의 내부를 탐구하기 위한 가장 잘 알려진 도구이다.

표현형 기반 스크리닝은 새로운 저분자 기반 신약 후보물질을 발굴하기 위한 중요한 출발점이다. 최근 수십 년간 많은 새로운 작용 기전을 지니는 신약들이 표현형 기반 스크리닝을 통해 발굴되었으며, 이는 표현형 기반 스크리닝이 신약개발을 위한 훌륭한 도구임을 시사한다. 하지만 얻어진 후보물질들의 자세한 메커니즘을 밝기 위해서는 추가적인 연구 과정과 노력이 필요하며, 이는 표현형 기반 신약개발에 있어서 피할 수 없는 과제이기도 하다. 적절한 화학적 도구의 개발은 표현형 기반 스크리닝의 초기

단계에서만 뿐만 아니라, 추후 표적 메커니즘 규명에 있어서도 중요한 단서를 제공해줄 수 있다.

표현형 기반 스크리닝을 출발점으로 하는 유용한 신약 개발 플랫폼은 다음과 같이 세 단계로 구성된다. (1) 표현형에 기반한 신약 후보 물질의 발굴, (2) 선택된 생리활성 저분자 화합물에 대한 표적 단백질의 규명, (3) 그리고 선택된 물질들의 생물학적 효능 개선을 위한 구조 최적화이다.

위의 플랫폼에 대하여 각 단계에서 적절한 화학적 도구들의 고안이 선행되어야 한다. 생명 현상의 표현형 관찰을 위한 형광체 및 형광 센서, 표적 규명을 위한 프로브, 개선된 생리활성 화합물을 구조활성 상관관계를 통하여 얻어내는 것까지, 각 단계에 필요한 도구들을 분자구조의 디자인을 통하여 고안할 수 있다. 본 학위논문에서는 위의 표현형 기반 신약 개발 플랫폼을 토대로, 생명 현상의 탐구를 위한 도구들인 바이오프로브와 생리활성 저분자 화합물을 디자인하고 최적화하는 연구에 관하여 이야기하고자 한다. 이를 다음 세 개의 파트로 나누어서 설명할 수 있겠다.

첫 번째 파트에서는 나프탈렌 구조에 기반한 이광자 형광체를 디자인하고 구축하는 연구에 대해 설명하고자 한다. 형광 유기 화합물은 생명체의 표현형 변화를 포함한 생명 현상을 형광을 통하여 관찰할 수 있게 해주는 유용한 도구이다. 나프탈렌 구조

기반 형광체인 아세단에서 출발하여, 가시광선 전 영역의 형광 물성을 나타내면서 손쉽게 합성 가능한 이광자 형광체들을 구축할 수 있었다. 나아가서, 과산화수소 농도를 감지하는 형광 장과장 영역의 이광자 형광 센서를 고안하였으며, 이를 체내 염증 반응의 모니터링에 적용할 수 있었다.

두 번째 파트에서는 생리활성물질의 표적 규명을 위한 프로브를 디자인하고 최적화한 연구에 대하여 설명하고자 한다. 첫 번째 섹션에서는 FITGE라는 방법에 기반한 표적 규명에 대하여 설명하고, 성공적인 두 가지 표적 규명 사례를 통하여 표적 규명 프로브 디자인 전략에 대하여 논하고자 한다. 두 번째 섹션에서는, 표적 규명 프로브에 사용되는 각 광반응성 작용기에 특이적으로 결합하는 단백질들을 찾아냄으로서, 표적 단백질 규명 효율을 개선한 연구에 대해 이야기하고자 한다. 세 번째 섹션에서는 표적 규명 프로브의 구조 변화, 특히 길이와 배향 변화에 따른 비특이적 단백질 결합 정도가 어떻게 바뀌게 되는지 논하고자 한다.

세 번째 파트는 생체 내 표적이 규명된 생리활성물질의 구조 최적화에 관한 연구이다. 첫 번째 섹션에서는, 천연물의 전합성과 그 유도체들의 합성을 통하여, ERR γ 의 역작용제들을 찾아내고 그 효능을 확인한 연구를 이야기하고자 한다. 두 번째 섹션에서는, 체내 염증 반응을 매개하는 HMGB1를 표적으로 하는 ICM 화합물을 구조 개선을 통하여 패혈증 치료에 성공적으로 적용한

연구를 이야기하고자 한다.

본 학위논문의 세 가지 파트에서 설명하고자 하는 바이오프로브와 생리활성 저분자 화합물에 대한 디자인 및 최적화에 관한 연구는 신규 신약 후보물질을 발굴하는 것뿐만 아니라, 잘 알려지지 않은 새로운 생명 현상을 탐구하는데 있어서 유용한 한 걸음이 될 것이라 기대하는 바이다.

주요어: 바이오프로브, 생리활성물질, 저분자 화합물, 합리적, 디자인, 최적화, 신약 개발, 표현형 기반 스크리닝, 형광체, 형광 프로브, 표적 규명, 광반응성 작용기, 분자 구조, ERR γ , 역작용제, 패혈증, HMGB1, 항염증, 구조-활성 상관관계

학번: 2010-23080