Supplementary Information

Development of bis-unsaturated ester aldehydes as amino-glue probes: Sequential double azaelectrocyclization as promising strategy for bioconjugation

Katsunori Tanaka^{*a}, Yuka Nakamoto^b, Eric R. O. Siwu^b, Ambara R. Pradipta^a, Koji Morimoto^c, Takeshi Fujiwara^c, Suguru Yoshida^d, Takamitsu Hosoya^d, Yuki Tamura^a, Go Hirai^a, Mikiko Sodeoka^a and Koichi Fukase^{*b}

^aRIKEN Advanced Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan, ^bDepartment of Chemistry, Graduate School of Science, Osaka University, 1-1 Machikaneyama-cho, Toyonaka-shi, Osaka 560-0043, Japan, ^cDivision of Biochemistry, Department of Melandar, Biochemistry, Curdwate School of

^cDivision of Biochemistry, Department of Molecular Biology and Biochemistry, Graduate School of Medicine/Faculty of Medicine, Osaka University, 1-3 Yamada-oka, Suita-shi, Osaka 565-0871, Japan,

^{*d}</sup>Laboratory of Chemical Bioscience, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-0062, Japan.*</sup>

e-mail: kotzenori@riken.jp (for KT) and koichi@chem.sci.osaka-u.ac.jp (for KF)

General Procedures. All commercially available reagents were used without further purification. Preparative separation was usually performed by column chromatography on silica gel (FUJI silysia LTD, BW-200 and BW-300) and by thin layer chromatography on silica gel (Merck, 20 x 20 cm, Silica gel 60 F_{254} , 1 mm). ¹H and ¹³C NMR spectra were recorded on both JEOL Lambda 500 NMR spectrometer and Delta ECA 500 NMR spectrometer, and chemical shifts were represented as δ -values relative to the internal standard TMS. MALDI-TOF-mass spectra were measured on PerSeptive Biosystems, Voyager RP-DE/H and SHIMADZU AXIMA-CFR mass spectrometers equipped with a nitrogen laser ($\lambda = 337$ nm). High resolution mass spectra (HRMS) were measured on a Bruker micro-TOF mass spectrometer under positive electrospray ionization (ESI) conditions. IR spectra were recorded on a *JASCO* FT/IR-8000 Fourier Transform Infrared Spectrometer. Fluorescence images of the labeled and/or bioconjugated cells were captured on OLYMPUS fluorescence microscope, IX71-23FL/DIC.



<u>Difficulty in dimerization</u>
1) Not stable under exposure of light
2) **R** = -SH: decomposition or isomerization of conjugated system through conjugate addition

3) $\mathbf{R} = -\mathbf{NH}_2$: self-condensation reaction through rapid azaelectrocyclization



Fig. SI-1 Dimerization trials of (*E*)-3-alkoxycarbonyl-5-phenyl-2,4-dienals.

Preparation of cis- and trans-3a. To a solution of sym-dibenzo-1,5-cyclooctadiene-3,7-diyne (2) (8.0 mg, 40 (mol) in dry CHCl₃ (3.0 mL) was added azide aldehyde (1a) (36.4 mg, 80 μ mol) in CHCl₃ (0.2 mL) at room temperature. After stirring for 2 h at this temperature, the mixture was concentrated in vacuo to yield the mixture of cis- and trans-3a in quantitative yield as a yellow solid (44.4 mg). The two isomers were separated by repeating the preparative thin layer chromatography on silica gel (CHCl₃ : MeOH = 25 : 1) to give one isomer of **3a** (isomer-1, 24 mg, 53%) and another isomer of **3a** (isomer-2, 11 mg, 25%). Data for **3a** (isomer-1): ¹H NMR (500 MHz, CDCl₃) $\delta = 10.07$ (d, 2H, J = 7.2 Hz), 9.23 (brs, NH x 2), 7.89 (brt, NH x 2), 7.63 (dd, 2H, J = 7.4, 1.4 Hz, 7.51 (d, 4H, J = 8.6 Hz), 7.42-7.49 (m, 6H), 7.38 (d, 4H, J = 8.6 Hz), 7.26 (d, 2H, J= 16.0 Hz), 6.98 (d, 2H, J = 16.0 Hz), 6.56 (d, 2H, J = 7.2 Hz), 4.37-4.41 (m, 2H), 4.27 (g, 4H, J = 7.2 Hz), 4.12-4.18 (m, 2H), 4.07 (t, 4H, J = 5.2 Hz), 2.00-2.14 (m, 4H), 1.56-1.61 (m, 2H), 1.33-1.43 (m, 6H), 1.30 (t, 6H, J = 7.2 Hz), 0.94-1.07 (m, 4H), 0.63-0.70 (m, 2H), 0.50-0.60 (m, 2H); ¹³C NMR (125 MHz, CDCl₃) δ =191.3 (2C), 175.0 (2C), 168.2 (2C), 166.5 (2C), 146.0 (2C), 144.9 (2C), 141.0 (2C), 139.4 (2C), 135.0 (2C), 132.3 (2C), 131.5 (2C), 131.2 (2C), 130.7 (2C), 130.4 (2C), 129.5 (2C), 129.3 (2C), 128.4 (4C), 126.5 (2C), 119.9 (4C), 118.2 (2C), 62.0 (2C), 48.5 (2C), 45.1 (2C), 35.8 (2C), 29.4 (2C), 28.0 (2C), 25.4 (2C), 25.1 (2C), 14.1 (2C); IR (KBr, cm⁻¹) 3276, 2931, 1663, 1530, 1246, 755; ESI-MS m/z calcd for $C_{62}H_{66}N_{10}O_{10}Na(M+Na)^+$ 1133.4856, found 1133.4857.

Data for **3a** (isomer-2): ¹H NMR (500 MHz, CDCl₃) $\delta = 10.06$ (d, 2H, J = 7.2 Hz), 9.37 (brs, N<u>H</u> x 2), 7.55-7.56 (m, 2H), 7.49-7.51 (m, 6H), 7.36-7.41 (m, 6H), 7.25-7.28 (m, 4H), 6.97 (d, 2H, J = 16.0 Hz), 6.56 (d, 2H, J = 7.2 Hz), 4.31-4.39 (m, 2H) 4.27 (q, 4H, J = 7.2 Hz), 4.09-4.15 (m, 2H), 3.92-4.02 (m, 4H), 2.11-2.15 (m, 4H), 1.64-1.70 (m, 4H), 1.45-1.50 (m, 4H), 1.29 (t, 6H, J = 7.2 Hz), 1.11-1.18 (m, 8H); ¹³C NMR (125 MHz, CDCl₃) $\delta = 191.4$ (2C), 174.6 (2C), 168.0 (2C), 166.4 (2C), 146.1 (2C), 145.9 (2C), 140.9 (2C), 139.1 (2C), 133.3 (2C), 132.3 (2C), 131.7 (2C), 130.8 (2C), 130.7 (2C), 130.4 (2C), 130.2 (2C), 129.2 (2C), 128.43 (4C), 128.35 (2C), 120.0 (4C), 118.4 (2C), 62.0 (2C), 48.5 (2C), 44.6 (2C), 35.6 (2C), 29.9 (2C), 28.4 (2C), 26.1 (2C), 25.1 (2C), 14.1 (2C); IR (KBr, cm⁻¹) 3280, 2931, 1663, 1520, 1247, 756; ESI-MS *m/z* calcd for C₆₂H₆₆N₁₀O₁₀Na (M+Na)⁺ 1133.4856, found 1133.4857.

Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry This journal is © The Royal Society of Chemistry 2013



Preparation of *cis-* **and** *trans-3b.* To a solution of aminoalcohol (structure shown below, 46 mg, 150 μ mol) in DMF (1.2 mL) was added azide-PEG4 NHS ester (58 mg, 150 μ mol) at room temperature. After stirring for overnight at this temperature, the mixture was concentrated *in vacuo*. The residue was purified by preparative thin layer chromatography on silica gel (CHCl₃ : MeOH = 8 : 1) to give the corresponding azide-PEG4 alcohol (54 mg, 62%) as a white solid: ¹H NMR (500 MHz, CDCl₃) δ = 8.84 (brs, N<u>H</u>), 7.60 (brs, N<u>H</u>), 7.54 (d, 2H, *J* = 8.6 Hz), 7.35 (d, 2H, *J* = 8.6 Hz), 6.82 (t, 1H, *J* = 6.0 Hz), 6.75 (d, 2H, *J* = 12.0 Hz), 4.54 (d, 2H, *J* = 6.0 Hz), 4.27 (q, 2H, *J* = 7.2 Hz), 4.08 (d, 2H, *J* = 6.0 Hz), 3.83 (t, 2H, *J* = 5.5 Hz), 3.70-3.72 (m, 2H), 3.60-3.64 (m, 12H), 3.37 (t, 2H, *J* = 5.2 Hz), 2.56 (t, 2H, *J* = 5.5 Hz), 1.34 (t, 3H, *J* = 7.2 Hz); ¹³C NMR (125 MHz, CDCl₃) δ = 172.9, 167.8, 166.9, 140.5 (2C), 138.0, 134.3, 132.8, 130.8, 127.3 (2C), 119.9 (2C), 70.6, 70.55 (2C), 70.5, 70.3, 70.2, 70.0, 67.5, 61.1, 59.8, 50.6, 44.5, 36.7, 14.3 ; IR (KBr, cm⁻¹) 3272, 2875, 2108, 1653, 1535, 1252, 1109; ESI-MS *m*/*z* calcd for C₂₇H₃₉N₅O₉Na (M+Na)⁺ 600.2640, found 600.2640.



To a solution of the azide-PEG4 alcohol obtained above (22 mg, 38 μ mol) in CH₂Cl₂ (2.0 mL) was added Dess-Martin periodinane (16 mg, 38 μ mol) at room temperature. After stirring for 15 min at this temperature, the mixture was concentrated under reduced pressure. The residue was purified by preparative thin layer chromatography on silica gel (CHCl₃: MeOH = 8 : 1) to give the corresponding aldehyde **1b** (16.5 mg, 76%) as a yellow solid: ¹H NMR (500 MHz, CDCl₃) δ = 10.14 (d, 1H, *J* = 7.2 Hz), 8.93 (brs, N<u>H</u>), 7.63 (d, 2H, N<u>H</u>, *J* = 8.6 Hz), 7.47 (d, 2H, *J* = 8.6 Hz), 7.32 (d, 1H, *J* = 16.0 Hz), 7.05 (d, 1H, *J* = 16.0 Hz), 6.64 (d, 1H, *J* = 7.2 Hz), 4.35 (q, 2H, *J* = 7.2

Hz), 4.10 (d, 2H, J = 6.0 Hz), 3.85 (t, 2H, J = 5.4 Hz), 3.72-3.74 (m, 2H), 3.60-3.65 (m, 12H), 3.38 (t, 2H, J = 5.0 Hz), 2.57 (t, 2H, J = 5.4 Hz), 1.38 (t, 3H, J = 7.2 Hz); ¹³C NMR (125 MHz, CDCl₃) $\delta = 191.3$, 172.9, 168.0, 166.5, 146.1, 141.1, 139.5, 131.4, 130.8, 128.4 (2C), 119.8 (2C), 118.1, 70.6, 70.5 (2C), 70.3 (2C), 70.2, 69.9, 67.6, 62.0, 50.6, 44.7, 36.7, 14.1 ; IR (KBr, cm⁻¹) 3317, 2872, 2104, 1668, 1532, 1248, 1114; ESI-MS *m*/*z* calcd for C₂₇H₃₇N₅O₉Na (M+Na)⁺ 598.2484, found 598.2484.



To a solution of *sym*-dibenzo-1,5-cyclooctadiene-3,7-diyne (**2**) (2.6 mg, 13 μ mol) in dry CHCl₃ (0.5 mL) was added aldehyde (**1b**) (15.1 mg, 26 μ mol) in CHCl₃ (1.0 mL) at room temperature. After stirring for 2 h at this temperature, the mixture was concentrated *in vacuo*. The residue was purified by column chromatography on silica gel (CHCl₃ : MeOH = 15 : 1) to give a mixture of *cis*- and *trans*-**3b** (17.7 mg, 100%) as a yellow solid: ¹H NMR (500 MHz, CDCl₃) δ = 10.14 (d, 2H, *J* = 7.2 Hz), 9.15 (brs, NH x 2), 7.39-7.72 (m, 18H), 7.32 (d, 2H, *J* = 15.8 Hz), 7.04 (d, 2H, *J* = 15.8 Hz), 6.63 (d, 2H, *J* = 6.9 Hz), 4.44-4.49 (m, 2H), 4.27-4.37 (m, 6H), 4.05-4.09 (m, 4H), 3.77-3.81 (m, 6H), 3.34-3.67 (m, 26H), 2.51-2.55 (m, 4H), 1.37 (t, 6H, *J* = 7.2 Hz); ¹³C NMR (125 MHz, CDCl₃) δ = 191.3 (2C), 172.8 (2C), 168.0 (2C), 166.5 (2C), 146.1 (2C), 145.6, 144.9, 141.2 (2C), 139.6 (2C), 135.4, 134.6, 132.5, 131.4 (2C), 131.31, 131.27, 130.9, 130.7 (2C), 130.4, 130.2, 130.1, 129.2, 129.1, 128.5, 128.4 (4C), 128.1, 126.4, 119.9 (4C), 118.0 (2C), 70.5 (2C), 70.5 (2C), 70.4 (2C), 70.34, 70.30, 70.21 (2C), 70.15 (2C), 69.2, 69.1, 67.4 (2C), 62.2, 62.0, 48.3, 48.2, 44.5 (2C), 36.7 (2C), 14.1 (2C); IR (KBr, cm⁻¹) 3324, 2924, 1668, 1532, 1250, 1113, 755; ESI-MS *m/z* calcd for C₇₀H₈₂N₁₀O₁₈Na (M+Na)⁺ 1373.5701, found 1373.5703.



General protocol for protein conjugation. To a solution of cyclic RGDyK peptide (620 μ g, 1.0

5

 μ mol) in PBS (25 μ L) was added a solution of *cis*- and/or *trans*-**3a** (550 μ g, 500 nmol) in DMSO (25 μ L) (concentration of **3a**: 10 mM). After the resulting solution was kept at 25 °C for 10 min, 10 μ L of the reaction mixture was added to a solution of HSA in PBS (66 μ g, 1.0 nmol, 90 μ L) (final reaction concentrations: 1.0×10^{-3} M for cyclic RGDyK+**3a** and 1.0×10^{-5} M for HSA, containing 5% DMSO). After the resulting solution was kept at 37 °C for 30 min, the mixture was directly analyzed by MALDI-TOF-MS.

Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry This journal is C The Royal Society of Chemistry 2013



Figs. SI-2 MALDI-TOF-MS spectra of mono-azaelectrocyclization products by the reaction of **3** with various peptides (peptide: 20 mM, **3**: 10 mM, 25 °C, 10 min in PBS buffer containing 50% DMSO).

(A) Cyclic RGDyK + **3a** (mixtures). Peak **a**: Cyclic RGDyK (m/z calcd for C₂₇H₄₂N₉O₈ (M+H)⁺ 620.3, found 620.5). Peak **b**: cyclic RGDyK + **3a** (mono-azaelectrocyclization product, m/z calcd for C₈₉H₁₀₆N₁₉O₁₇ (M-O+H)⁺ 1696.8, found 1696.9). Peak **c**: 2 molecules of RGDyK + **3a** (homo-coupling product, not detected). Peak **d**: **3a** (not detected).

(B) Cyclic RGDyK + **3b** (mixtures). Peak **a**: Cyclic RGDyK. Peak **b**: Cyclic RGDyK + **3b** (mono-azaelectrocyclization product, m/z calcd for C₉₇H₁₂₂N₁₉O₂₅ (M-O)⁺ 1936.9, found 1936.6). Peak **c**: 2 molecules of RGDyK + **3b** (homo-coupling product, not detected). Peak **d**: **3b** (not detected).

(C) Biotin-PEG-NH₂ + **3a** (mixtures). Peak **a**: Biotin-PEG-NH₂ (m/z calcd for C₁₈H₃₅N₄O₅S (M+H)⁺ 419.2, found 419.8). Peak **b**: Biotin-PEG-NH₂+**3a** (mono-azaelectrocyclization product, m/z calcd for C₈₀H₉₉N₁₄O₁₄S (M-O)⁺ 1494.7, found 1494.8). Peak **c**: 2 molecules of biotin+**3a** (homo-coupling product, almost negligible). Peak **d**: **3a** (almost negligible).

(D) Somatostatin + **3a** (mixtures). Peak **a**: Somatostatin (m/z calcd for C₇₆H₁₀₄N₁₇O₂₀S₂ (M+H)⁺ 1638.7 found 1638.7). Peak **b**: Somatostatin + **3a** (mono-azaelectrocyclization product, m/z calcd for C₁₃₈H₁₆₈N₂₇O₂₉ (M+H-H₂O)⁺ 2714.2, found 2714.3). Peak **c**: 2 molecules of somatostatin + **3a** (homo-coupling product, very small peak).

(E) Disialoglycopeptide $4\mathbf{a} + 3\mathbf{a}$ (mixtures). Peak **a**: Disialoglycopeptide $4\mathbf{a}$ (*m/z* calcd for $C_{112}H_{190}N_{15}O_{70}$ (M+H)⁺ 2866.2 found 2866.1). Peak **b**: $4\mathbf{a} + 3\mathbf{a}$ (mono-azaelectrocyclization product, *m/z* calcd for $C_{174}H_{254}N_{25}O_{79}$ (M-O+H)⁺ 3942.7, found 3942.4). Peak **c**: 2 molecules of $4\mathbf{a} + 3\mathbf{a}$ (not detected).



Figs. SI-3 MALDI-TOF-MS of HSA (10^{-5} M) conjugated to the biotin-PEG-NH₂ pretreated with **3a** (**3a**: 10^{-3} M, 37 °C, 30 min in PBS buffer containing 5% DMSO). (a) Intact HSA. (b) Conjugation in the presence of the *cis*- and *trans*-mixture of **3a** (theoretical mass unit difference for each conjugation product: m/z = 1493.8). (c) Conjugation of one isomer of **3a** (isomer-1). (d) Conjugation of another isomer of **3a** (isomer-2).



Fig. SI-4 MALDI-TOF-MS spectrum of the reaction between bis(NHS)PEG₅ and cyclic RGDyK under the identical conditions applied to Figs. SI-2. Structure of bis(NHS)PEG₅ is shown in Fig. SI-1 (X = $-CH_2OCH_2$ -, n = 5). *Note that the significant amount of the homo-coupling product is produced*. RGDyK+bis(NHS)PEG₅ (mono-coupling product, *m/z* calcd for C₄₅H₆₉N₁₀O₁₈ (M+H)⁺ 1037.5, found 1038.3). 2 Molecules of RGDyK+bis(NHS)PEG₅ (homo-coupling product, *m/z* calcd for C₆₈H₁₀₅N₁₈O₂₃ (M+H)⁺ 1541.8, found 1542.7).



Fig. SI-5 MALDI-TOF-MS of HSA (10^{-5} M) conjugated to the mono-coupling product of RGDyK peptide with bis(NHS)PEG₅ (10^{-3} M), which was obtained in **Fig. SI-4** and isolated by HPLC. Bioconjugation was performed under the identical conditions in Fig. 3. Theoretical mass unit difference for each conjugation product: m/z = 923.5.

FITC-labeling of disialoglycopeptide 4a. Disialoglycopeptide **4a** (3.0 mg, 1.0 μ mol) was added to a solution of 5(6)-carboxyfluorescein *N*-hydroxysuccinimide ester (470 μ g, 1.0 μ mol) in DMF (200 μ L) and H₂O (200 μ L) at room temperature. After the mixture was stirred overnight at room temperature, the solvent was removed *in vacuo* and the residue was purified by HPLC [column: Nacalai Tesque 5C₁₈-AR300, 4.6×250 mm; MeCN in H₂O (containing 0.1% TFA); 10-100% gradient over 40 min; 1 mL/min; UV detection at 245 nm]. The fractions **a** (retention time at 11.9 min), **b** (retention time at 12.5 min), and **c** (retention time at 12.8 min) contain the desired mono-labeled products, which were lyophilized to give the FITC-labeled **4b**, 930 μ g, 770 μ g, and 780 μ g, respectively. Fraction **a**, which was expected to be the *N*-terminus labeling product, was used for cell surface bioconjugation trials: MALDI-TOF-MS detected the three peaks characteristic for disialoglycans (see ref. 8c and 14 in the text): *m/z* calcd for C₁₃₃H₂₀₀N₁₅O₇₆ (disialoside-form, M+H)⁺ 3225.1, found 3225.6, for C₁₂₂H₁₈₁N₁₄O₆₈Na (monosialoside-form, M-H+Na)⁺ 2954.8, found 2954.3, and for C₁₁₁H₁₆₄N₁₃O₆₀Na₂ (asialo-form, M-H+Na+Na)⁺ 2686.5, found 2686.5.



Fig. SI-6 Reaction of disialoglycopeptide with FITC-NHS ester and HPLC profile.

Cell culture and cell surface conjugation. HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) and 5% antibyomyco. Cells of 2 x 10^5 were seeded on 13 mm diameter glass cover slips and grown overnight at 37 °C, 5% CO₂. Cells were washed twice with PBS (pH 7.2-7.4) to exclude residual conjugation inhibitors. Washing buffer was replaced with either TAMRA-NH₂+**3a** or the disialoglycopeptide **4b**+**3a** in PBS at 1 x 10^{-6} M concentration containing 0.005% DMSO, which were prepared according the procedure described above. After incubation for 5 min at 37° C, 5% CO₂, cells were washed twice with the PI (propidium iodide) (Invitrogen, 3.5 μ g/700 μ L in medium) for 30 min at 37° C, 5% CO₂. Cells were then fixed with 2% paraformaldehyde and 4% sucrose in PBS for 15 min at room temperature. After nuclei and chromosomes were labeled with 2 μ g/mL DAPI (4',6-diamino-2-phenylindole, Invitrogen), they were mounted with Prolong Gold (Invitrogen) to analyze by microscopy.



Fig. SI-7 Microscopy images of HeLa cells treated with disialoside 4b+3a at a concentration of 10^{-4} M in PBS for 5 min at 37 °C, 5% CO₂. Left panel: phase contrast, right panel: FITC.



Fig. SI-8 Microscopy images of HeLa cells treated with medium, PBS, **4b+3a**, and staurosporine for 5 min at 37 °C, 5% CO₂; bars indicate 100 μ m. Cells treated with (a) medium, (b) PBS, (c) 1 x 10⁻⁶ M of **4b+3a** (a mixture of the *cis-* and *trans-*isomers), and (d) 2 μ M of staurosporine to induce the apoptosis. Cells were treated with PI (red) to check viability. Nuclei and chromosomes were also labeled with DAPI (blue). From left to right: phase contrast, DAPI, and PI.



Fig. SI-9 ¹H NMR spectrum of one isomer of **3a** (isomer-1).



Fig. SI-10 ¹³C NMR spectrum of one isomer of **3a** (isomer-1).



Fig. SI-11 ¹H NMR spectrum of another isomer of **3a** (isomer-2).



Fig. SI-12 ¹³C NMR spectrum of another isomer of **3a** (isomer-2).





Fig. SI-13 ¹H NMR spectrum of azide-PEG4 alcohol.





Fig. SI-14 ¹³C NMR spectrum of azide-PEG4 alcohol.





Fig. SI-15 ¹H NMR spectrum of **1b**.





Fig. SI-16 ¹³C NMR spectrum of 1b.





Fig. SI-17 ¹H NMR spectrum of *cis*- and *trans*-**3b** as mixtures.





Fig. SI-18 ¹³C NMR spectrum of *cis*- and *trans*-**3b** as mixtures.