Electronic Supplemental Information

Detecting ligand-protein interactions inside cells using reactive peptide tags and split luciferase

Tsuyoshi Takahashi[a] and Hiroaki Hagiwara[a]

[a]Graduate School of Science and Technology, Gunma University, 1-5-1, Tenjin-cho, Kiryu, Gunma, 376-8515, Japan

Contents

Experimental section ---------- 2
Figure S1 ---------- 22
Figure S2 ---------- 23
Figure S3 ---------- 24
Figure S4 ---------- 25
Figure S5 ---------- 26
Figure S6 ---------- 27
Figure S7 ---------- 28
Figure S8 ---------- 29
Figure S9 ---------- 30
Figure S10 ---------- 31
Figure S11 ---------- 32

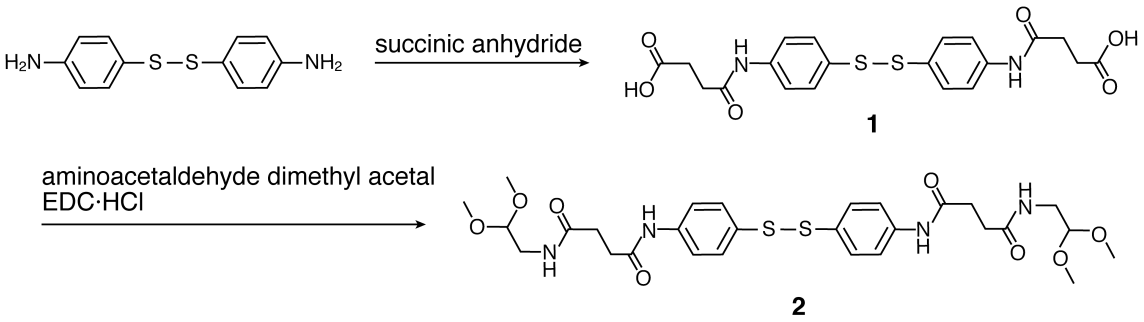
Experimental Section

Materials and methods

All chemicals and solvents were of reagent or HPLC grade. Amino acid derivatives and reagents for peptide synthesis were purchased from Watanabe Chemical and AAPPTEC. Other chemicals and reagents were purchased from Tokyo Chemical Industry, Nacalai Tesque, Sigma Aldrich, TOYOBO, and QIAGEN. NMR spectra were recorded on a JEOL JNM-ECX400 spectrometer (400 MHz). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOFMS) was measured on a Shimadzu AXIMA Performance mass spectrometer using 2-cyano-3-(4-hydroxyphenyl)acrylic acid as a matrix. Electrospray ionization mass spectrometry was measured on a Shimadzu LCMS-2020 single quadrupole liquid chromatography mass spectrometer. HPLC was performed on COSMOSIL 5C18-AR II (4.6 x 150 mm and 10 x 250 mm, Nacalai Tesque) and Daisopak SP-120 C4-Bio (10 x 250 mm, Daiso Chemical) packed columns by employing Hitachi L-7100 and Shimadzu SCL-10Avp HPLC systems. PCR was performed on a MiniCycler PTC-150 thermal cycler (MJ Research). DNA sequence was analyzed by a 3130 Genetic Analyzer (Applied Biosystems) using BigDye Terminator 3.1 cycle sequencing kit. UV-visible absorption spectra were recorded on a Shimadzu UV-2450 spectrophotometer using a quartz cell with 1.0 cm or 0.5 cm pathlength. Histidine-tagged proteins were purified on a Ni-NTA agarose resin (QIAGEN). Buffer solution was exchanged by a PD10 desalting column (GE Healthcare).

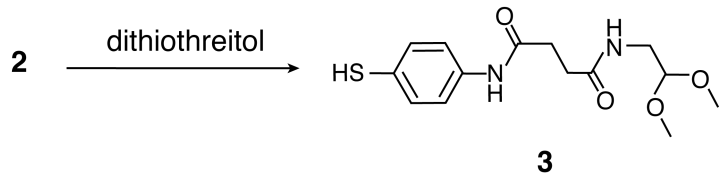
Synthesis of NLcn-appended ligands

Synthesis of N^1,$N^{1'}$-(disulfanediylbis(4,1-phenylene))bis(N^4-(2,2-dimethoxy ethyl)succinamide) (compound **2**):



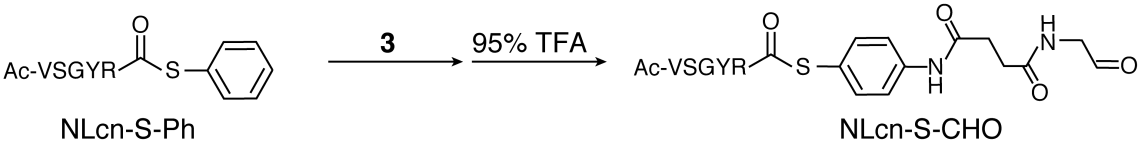
To a solution of 4,4'-dithiodianiline (1.0 g, 4.03 mmol) in toluene (25 mL) was added succinic anhydride (0.805 g, 8.05 mmol). The mixture was heated for 2 hours under reflux, and then cooled to room temperature. Hexane was added and the resulting solid was collected by filtration. Half amounts of the crude product of compound **1** (0.804 g, ≈ 1.7 mmol) and aminoacetaldehyde dimethylacetal (0.585 mL, 5.37 mmol) in dimethylformamide (DMF; 10 mL) was added 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC·HCl; 1.37 g, 7.15 mmol) at 4°C. The mixture was stirred for 6 hours at 4°C. The solvent was removed by evaporation and water was added. The resulting solid was washed with water and collected by filtration to afford compound **2** (0.933 g, 1.58 mmol, 88% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 10.11 (s, 2H), 7.99 (t, J = 5.7 Hz, 2H), 7.59 (d, J = 8.7 Hz, 4H), 7.41 (d, J = 8.7 Hz, 4H), 4.31 (t, J = 5.5 Hz, 2H), 3.24 (s, 12H), 3.14 (t, J = 5.7 Hz, 4H), 2.53 (t, J = 6.9 Hz, 4H), 2.41 (t, J = 6.9 Hz, 4H).

Synthesis of N^1-(2,2-dimethoxyethyl)-N^4-(4-mercaptophenyl)succinamide (compound **3**):



To a solution of compound **2** (0.40 g, 0.677 mmol) in DMF (20 mL) was added dithiothreitol (DTT; 316 mg, 2.05 mmol). The mixture was stirred for 6 hours at room temperature. The solvent was removed by evaporation and water was added. The resulting solid was washed with water and collected by filtration to afford compound 3 (0.23 g, 0.73 mmol, 54% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.93 (s, 1H), 7.99 (t, J = 5.9 Hz, 1H), 7.47 (d, J = 8.7 Hz, 2H), 7.20 (d, J = 8.7 Hz, 2H), 5.26 (s, 1H), 4.31 (t, J = 5.5 Hz, 1H), 3.25 (s, 6H), 3.14, (t, J = 5.7 Hz, 2H), 2.51 (t, J = 6.9 Hz, 2H), 2.41 (t, J = 6.9 Hz, 2H).

Synthesis of NLcn-CHO:

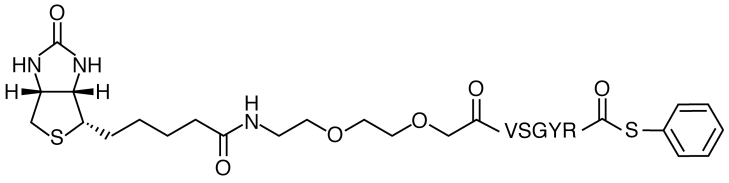


To synthesize NLcn-CHO, a peptide thioester bearing the thiophenyl group was firstly prepared by the thioester exchange reaction of peptide *N*-acyl-benzimidazolinone (Nbz).[1] The protected peptide was elongated on Rink amide MBHA resin (AAPPTEC). 3-Fmoc-4-diaminobenzoic acid (Dbz) was synthesized according to the reported procedure.[1] Fmoc-Dbz-OH (164 μmol) was coupled on the resin (41 μmol) by use of *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) (328 μmol) and

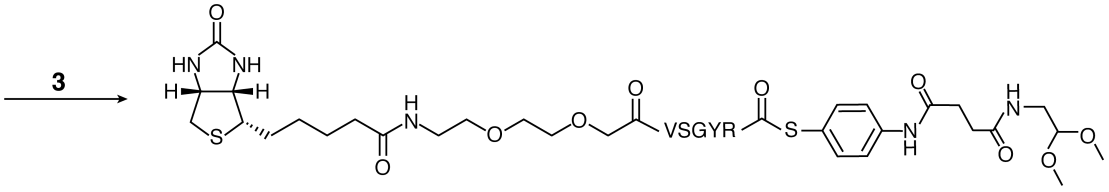
diisopropyl ethyl amine (DIEA) (296 μmol) in 1-methyl-2-pyrrolidone (NMP). After removal of Fmoc group by 20% (v/v) piperidine in NMP, Fmoc-Arg(Pbf)-OH was coupled on the resin with HATU (164 μmol) and DIEA (296 μmol). After that, peptide elongation was performed by use of Fmoc-AA-OH (135 μmol), *O*-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphoate (HBTU) (135 μmol), 1-hydroxybenzotriazole hydrate ($HOBt{\cdot}H_2O$) (135 μmol), and DIEA (202.5 μmol). After the elongation of Fmoc-AA-OH, acetic acid (135 mmol) was coupled on the peptidyl resin with HBTU (135 μmol), $HOBt{\cdot}H_2O$ (135 μmol), and DIEA (202.5 μmol). The resin was then washed with dichloromethane (DCM) and *p*-nitrochloroformate (225 μmol) in DCM was added and mixed for 45 min. Then, the resin was washed with DCM, and treated with 0.5 M DIEA in NMP (2 mL) for 30 min. After washing with NMP and DCM, the peptidyl-resin was dried under vacuum and treated with trifluoroacetic acid (TFA) in the presence of 2.5% (v/v) water and 2.5% (v/v) triisopropyl silane (TIS) for 2 h. The reaction solution was concentrated by evaporation, and the residue was poured into cold diethyl ether. The resulting precipitate was collected by centrifugation and dried under vacuum. The crude peptide was dissolved in water (2 mL) and acetonitrile (2 mL), and thiophenol (150 μL) was added. 0.1 m sodium bicarbonate (pH 8.0) (1 mL) was added. Afer 2.5 h, TFA (0.1 mL) was added and thiophenol was removed by extraction with diethyl ether. The resulting solution was lyophilized. Crude NLcn-S-Ph (≈ 19 μmol) and compound **3** (38 μmol) were dissolved in water (500 μL) and acetonitrile (500 μL). 0.1 M Sodium bicarbonate (500 μL) was added. After completion of the reaction (≈ 2 h, checked by HPLC), TFA (20 mL) was added and the mixture was stirred for 10 min. Water was added, and the solvent was removed by evaporation. The residue was purified on a semi-preparative column (Cosmosil, AR-II) to afford NLcn-CHO (1.77 mg,

1.80 μmol, 9.3% yield). The purified NLcn-CHO was lyophilized and identified by ESI-MS; *m*/*z* calcd for $C_{39}H_{55}N_{10}O_{11}S$ (exact mass) $[M+H]^+$ 871.38, found 871.14.

Synthesis of Bio-NLcn-CH(OMe)$_2$:



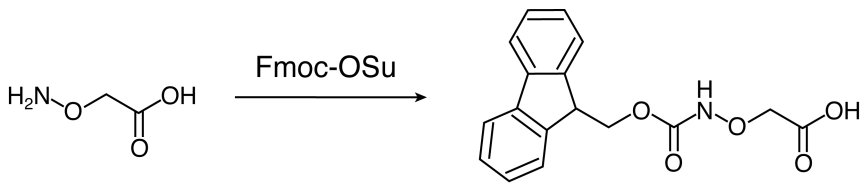
Bio-NLcn-S-Ph



Bio-NLcn-CH(OMe)$_2$

Bio-NLcn-S-Ph was synthesized by a similar procedure for NLcn-S-Ph. Crude Bio-NLcn-S-Ph (≈ 16 μmol) and compound **3** (36 μmol) were dissolved in water (1.5 mL) and tifluoroethanol (1.5 mL). 1 M Tris·HCl (pH 8.0) (100 μL) was added. After completion of the reaction (≈ 3.5 h, checked by HPLC), the product was purifed on a semi-preparative column (Daisopack, SP-120 C4-BIO) to afford Bio-NLcn-CH(OMe)$_2$ (4.54 mg, 3.64 μmol, 18% yield). The purified Bio-NLcn-CH(OMe)$_2$ was lyophilized and identified by ESI-MS; *m*/*z* calcd for $C_{55}H_{84}N_{13}O_{16}S_2$ (exact mass) $[M+H]^+$ 1246.56, found 1246.69.

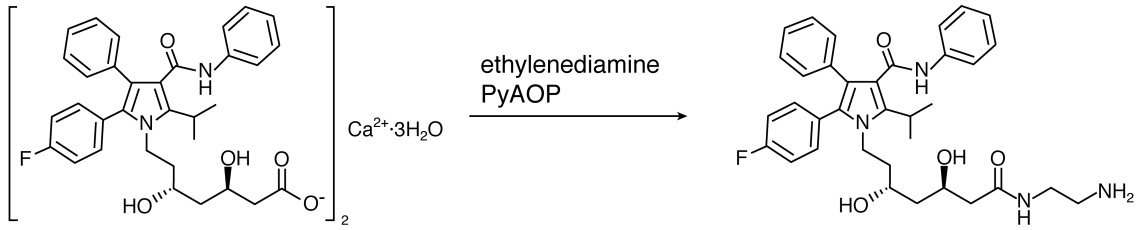
Synthesis of Fmoc-aminooxy acetic acid:



Fmoc-aminooxyacetic acid

To a solution of carboxymethoxylamine hemihydrochloride (1.0 g, 9.15 mmol) and triethylamine (2.56 mL, 18.3 mmol) in water (10 mL) was added *N*-[(9*H*-fluoren-9-ylmethoxy)carbonyloxy]succinimide (Fmoc-OSu) (2.78g, 8.23 mmol) in 1,4-dioxane (10 mL). The reaction was monitored by TLC. After 2 h, 10% citric acid aqueous solution (5 mL) was added. The solvent was removed by evaporation and the residue was extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, and concentrated by reduced pressure. Hexane was added and the resulting solid was collected to afford Fmoc-aminooxy acetic acid (2.02 g, 6.43 mmol, 78% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 7.89 (d, *J* = 7.3 Hz, 2H), 7.69 (d, *J* = 7.3 Hz, 2H), 7.44 – 7.30 (m, 4H), 4.36 (d, *J* = 6.9 Hz, 2H), 4.26-4.23 (m, 3H).

Synthesis of atorvastatin analog **1**:

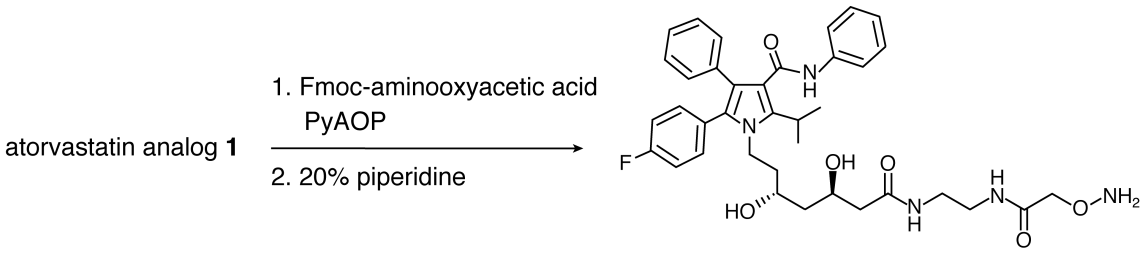


atorvastatin calcium salt trihydrate — atorvastatin analog **1**

To a solution of atorvastatin calcium salt trihydrate (purchased from TCI) (100 mg, 82.7 μmol) and ethylenediamine (54 μL, 0.866 mmol) in DMF (0.5 mL) was added (7-azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphoate (PyAOP) (110 mg, 211 μmol). After 24 h, the solvents were removed by evaporation and the residue was purified by silica gel chromatography with $CHCl_3$: MeOH : NH_4OH = 5 : 1 : 0.1 to afford atorvastatin (80 mg, 133 μmol, 77% yield). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.81 (s, 1H), 7.75 (t, J = 5.5 Hz, 1H), 7.50 (d, J = 7.3 Hz, 2H), 7.26-7.16 (m, 6H),

7.06 (m, 4H), 7.02-7.96 (m, 2H), 4.65 (br, s, 1H), 3.94 (m, 1H), 3.82 (m, 1H), 3.74 (m, 1H), 3.60-3.48 (m, 3H), 3.22 (sept, J = 7.3 Hz, 1H), 3.02 (q, J = 6.1 Hz, 2H), 2.53 (t, J = 6.4 Hz, 2H), 2.12 (d, J = 6.4 Hz, 2H) 1.68-1.48 (m, 2H), 1.40 (m, 1H), 1.37 (d, J = 7.3 Hz, 6H), 1.29 (m, 1H).

Synthesis of atorvastatin analog **2**:

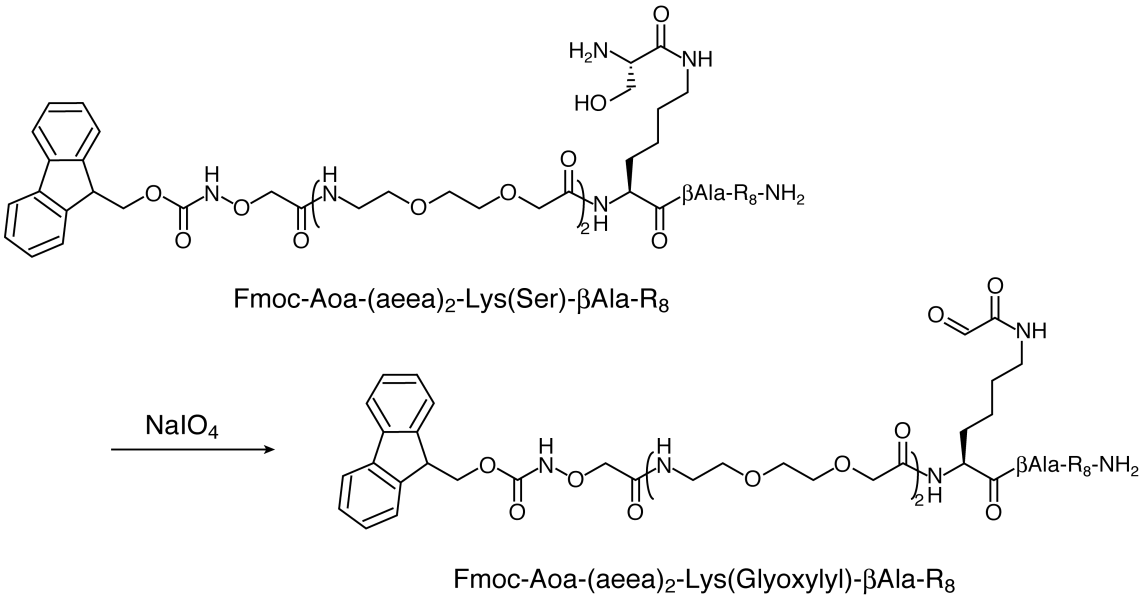


atorvastatin analog **2**

To a solution of Fmoc-aminooxyacetic acid (67 mg, 222 μmol), atorvastatin analog **1** (89 mg, 148 μmol), and PyAOP (93 mg, 178 μmol) in DMF (0.5 mL) was added triethylamine (41 μL, 293 μmol). The reaction mixture was stirred for 2.5 h. ADMF (1 mL) and piperidine (0.375 mL) was then added. After completion of the deprotection of Fmoc group, the mixture was concentrated by reduced pressure and the residue was purified by silica gel column chromatography with $CHCl_3$: MeOH : NH_4OH = 9 : 1 : 0.1 to give atorvastatin analog **2** (66 mg, 98 μmol, 66% yield). ^1H NMR (400 MHz, DMSO-d_6): δ = 9.81 (s, 1H), 7.90 (t, J = 5.0 Hz, 1H), 7.82 (t, J = 5.0 Hz, 1H), 7.50 (d, J = 7.8 Hz, 2H), 7.26-7.16 (m, 6H), 7.06 (m, 4H), 7.02-6.96 (m, 2H), 6.35 (s, 1H), 4.74 (d, J = 4.6 Hz, 1H), 4.63 (d, J = 5.0 Hz, 1H), 3.98-3.90 (m, 1H), 3.92 (s, 2H), 3.83 (m, 1H), 3.74 (m, 1H), 3.52 (s, 1H), 3.22 (sept, J = 6.9 Hz, 1H), 3.14 (m, 4H), 2.12 (d, J = 6.4 Hz, 2H), 1.63 (m, 1H), 1.53 (m, 1H), 1.40 (m, 1H), 1.37 (d, 6.9 Hz, 6H), 1.29 (m, 1H).

Synthesis of Fmoc-Aoa-$(aeea)_2$-Lys(Glyoxylyl)-βAla-R_8:

Synthesis of the peptide bearing serine upon the lysine side-chain and

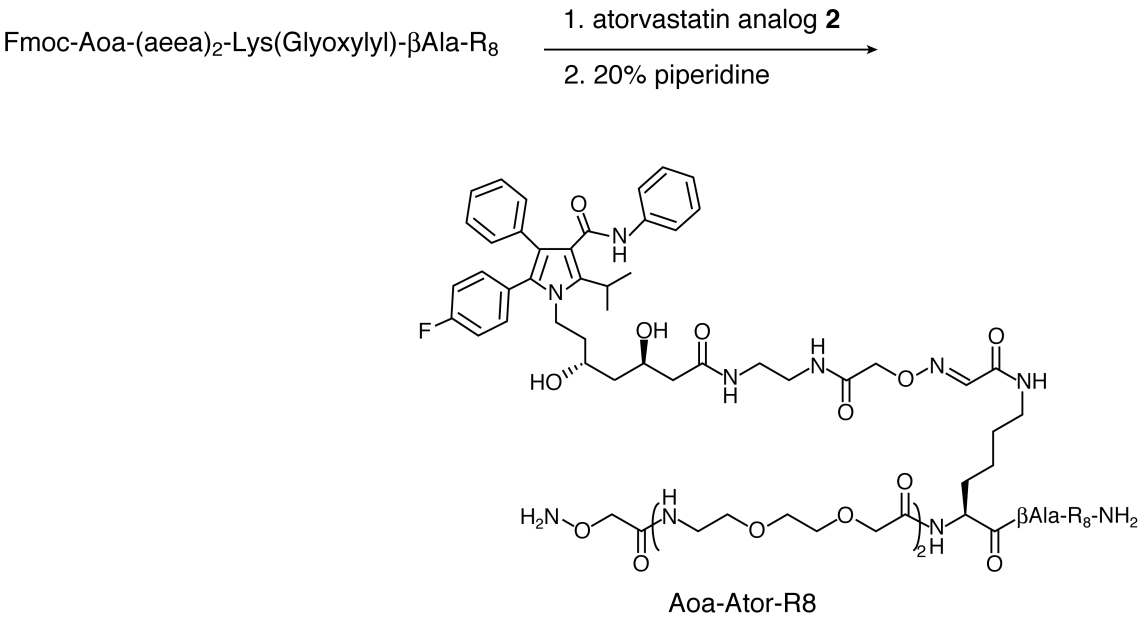


octaarginine was performed on a Rink Amide MBHA resin (AAPPTEC) by employing Fmoc-Lys(Mtt)-OH. After elongation of the peptide main-chain, Fmoc-aminooxyacetic acid N-succimidyl ester (6 equiv) in NMP was added to the resin, and the reaction tube was rotated for 16 h at room temperature. After washing with NMP and DCM, the peptidyl-resin was treated with 1% (v/v) TFA and 5% (v/v) TIS in DCM (1 min x 20). After removal of Mtt, the resin was washed with DCM and NMP, and neutralized with 5% (v/v) DIEA in NMP. Boc-Ser(tBu)-OH (3 equiv) was coupled to the Lys side-chain using HATU (3 equiv) and DIEA (6 equiv) for 45 min. After washing with NMP and DCM, the peptidyl-resin was dried under vacuum and treated with TFA as described above to give crude peptide, Fmoc-Aoa-$(aeea)_2$-Lys(Ser)-βAla-R_8 (37 mg, 12 μmol, 48%). The peptide was identified MALDI-TOFMS; m/z calcd for $C_{89}H_{156}N_{40}O_{22}$ (exact mass) $[M+H]^+$ 2137.23, found 2137.56.

The crude peptide was dissolved in 6 mL of 0.2 M phosphate (pH 7.2),

and then 100 μL of 0.3 M $NaIO_4$ in water was added, and the reaction mixture was stirred for 1 h at room temperature. The product was purified on a semi-preparative column (Cosmosil, AR-II) to afford Fmoc-Aoa-$(aeea)_2$-Lys(Glyoxylyl)-βAla-R_8 (11.4 mg, 3.8 μmol, 15% yield). The purified peptide was lyophilized and identified by MALDI-TOFMS; *m*/*z* calcd for $C_{88}H_{151}N_{39}O_{22}$ (exact mass) $[M+H]^+$ 2106.19, found 2106.62.

Synthesis of Aoa-Ator-R8:



Aoa-Ator-R8

Fmoc-Aoa-$(aeea)_2$-Lys(Glyoxylyl)-βAla-R_8 (5.6 mg, 1.9 μmol) was dissolved in 3% acetic acid in water (1.5 mL), and then 500 μL of atorvastatin analog 2 (1.8 mg, 2.7 μmol) in acetonitrile was added. The reaction mixture was stirred for 2 days at room temperature. After completion of the reaction checked by RP-HPLC, the solvent was removed by evaporation, and 20% (v/v) piperidine in DMF was added. After 30 min, ice-cold diethylether was added to precipitate the product. After centrifugation, the solid was washed twice with

diethylether and dissolved in 30% acetic acid in water. The product was purified on a semi-preparative column (Daisopack, SP-300 C4-BIO) to afford Aoa-Ator-R_8 (2.0 mg, 0.59 μmol, 22% yield). The purified product was identified by MALDI-TOFMS; *m*/*z* calcd for $C_{110}H_{184}N_{44}O_{25}F$ (exact mass) $[M+H]^+$ 2540.45, found 2540.75.

Synthesis of NLcn-appended ligands and acetylated peptides:

Synthesis of the peptides bearing (aminooxy)acetyl group (Aoa-Peps) were performed on a wang-PEG resin (Watanabe Chemical Industry) or a Rink amide MBHA resin (AAPPTEC). Whereas β-alannine was used as a linker between the binding sequence and (aminooxy)acetyl group to synthesize NLcn-pYEEI, NLcn-YEEI, and NLcn-Dpep1n, two 2-[2-(2-Aminoethoxy)ethoxy]acetic acid residues were used as a linker to synthesize NLcn-Dpep1, NLcn-Dpep2, NLcn-Lpep1, NLcn-cdc25c-T48, and NLcn-PMI. To synthesize After elongation of Fmoc-AA-OH on the resin, [(*tert*-butoxycarbonyl)aminooxy] acetic acid *N*-succinimidyl ester (Boc-Aoa-OSu) (10 equiv) in NMP was added to the resin, and the reaction tube was rotated for overnight at room temperature. The peptidyl-resin was treated with TFA and Aoa-Peps were purified on a semi-preparative column as described above.

A series of Aoa-Peps, atorvastatin analog **2**, and Aoa-Ator-R8 were reacted with NLcn-CHO by the following procedure: 10 μL of 10 mM Aoa-Pep or atorvastatin derivative in DMSO and 10 μL of 10 mM NLcn-CHO in DMSO were mixed with 20 μL of 20 mM acetic acid aqueous solution, and then the reaction mixture was incubated for 2 h at room temperature. The reaction was monitored by RP-HPLC. After completion of the reaction, the solution was stored at -80°C. Biotin-appended peptides were prepared by the following

procedure: 10 μL of 10 mM Aoa-Pep in DMSO and 10.9 μL of 9.16 mM Bio-NLcn-CH(OMe)2 were mixed and lyophilized, and then the powder was dissolved in 20 μL of 95% (v/v) aqueous TFA. After 5 min, 400 μL of water was added and lyophilized. The reaction completion was checked by RP-HPLC after dissolving in DMSO.

Products except for NLcn-Dpep1n were identified by ESI-MS;

NLcn-pYEEI, *m/z* calcd for $C_{69}H_{98}N_{16}O_{26}PS$ $[M+H]^+$ 1630.65, found 1630.33.

NLcn-YEEI, *m/z* calcd for $C_{69}H_{97}N_{16}O_{23}S$ $[M+H]^+$ 1550.67, found 1550.51.

Bio-NLcn-pYEEI, *m/z* calcd for $C_{83}H_{122}N_{19}O_{30}PS_2$ $[M+2H]^{2+}$ 980.54, found 980.20.

Bio-NLcn-YEEI, *m/z* calcd for $C_{83}H_{120}N_{19}O_{27}S_2$ $[M+H]^+$ 1880.08, found 1879.54.

NLcn-Ator, *m/z* calcd for $C_{76}H_{97}FN_{15}O_{16}FS$ $[M+H]^+$ 1527.74, found 1527.36.

NLcn-Dpep1, *m/z* calcd for $C_{147}H_{236}N_{55}O_{39}S$ $[M+3H]^{3+}$ 1154.62, found 1154.20.

NLcn-Dpep2, *m/z* calcd for $C_{149}H_{236}N_{55}O_{39}S_2$ $[M+3H]^{3+}$ 1173.31, found 1173.32.

NLcn-Lpep1, *m/z* calcd for $C_{147}H_{236}N_{55}O_{39}S$ $[M+3H]^{3+}$ 1154.62, found 1154.20.

NLcn-cdc25c-T48, *m/z* calcd for $C_{140}H_{246}N_{58}O_{39}PS$ $[M+3H]^{3+}$ 1142.95, found 1142.48.

NLcn-PMI, *m/z* calcd for $C_{178}H_{287}N_{63}O_{48}S$ $[M+4H]^{4+}$ 1027.41, found 1027.42.

NLcn-Ator-R8, *m/z* calcd for $C_{149}H_{239}FN_{54}O_{35}S$ $[M+4H]^{4+}$ 849.07, found 849.55.

NLcn-LpepΔ, *m/z* calcd for $C_{147}H_{235}N_{55}O_{36}S$ $[M+4H]^{4+}$ 846.22, found 846.13.

Bio-NLcn-Dpep2, *m/z* calcd for $C_{163}H_{264}N_{58}O_{43}S_3P$ $[M+5H]^{5+}$ 770.27, found 770.20.

Ac-Dpep1n, *m/z* calcd for $C_{39}H_{51}N_7O_{11}P$ (exact mass) $[M+H]^+$ 824.34, found 824.35.

Ac-Dpep2n, *m/z* calcd for $C_{41}H_{51}N_7O_{11}SP$ (exact mass) $[M+H]^+$ 880.92, found 880.20.

Ac-Dpep1, *m/z* calcd for $C_{96}H_{164}N_{42}O_{22}P$ $[M+3H]^{3+}$ 763.19, found 763.01.

NLcn-Dpep1n was identified by MALDI-TOFMS; *m*/*z* calcd for $C_{81}H_{108}N_{19}O_{23}SP$ (exact mass) $[M+H]^+$ 1778.74, found 1778.33.

Expression of NanoLuc 11S

The plasmid encoding NanoLuc 11S, termed pET28-11S, was prepared as described previously.[2] The plasmid pET28-11S was transformed into BL21(DE3). The transformed *E. coli* cells were grown in LB broth at 37°C in the presence of kanamycin (34 μg/mL) and chloramphenicol (34 μg/mL) until OD_{600} = 0.5. After that the flask was kept at 18°C for 30 min and then induced by 0.4 mM IPTG to express at 18°C for 16 h. Cells were collected by centrifugation and frozen at -80°C. Cells were resuspended in a bind buffer [50 mM phosphate (pH 8.0), 0.5 M NaCl, 5 mM imidazole], and sonicated on ice. After centrifugation (10,000 g for 70 min at 4°C), the supernatant was loaded onto Ni-NTA agarose resin. After washing with a wash buffer [50 mM phosphate (pH 8.0), 0.5 M NaCl, 30 mM imidazole], the protein was eluted with an elution buffer [50 mM phosphate (pH 8.0), 0.5 M NaCl, 0.5 M imidazole]. The buffer was exchanged by PD10 desalting column into storage buffer [50 mM Tris·HCl (pH 7.0), 0.3 M NaCl, 1 mM EDTA, 2 mM DTT, 10% (v/v) glycerol].

Preparation of NLcc-appended proteins

The genes encoding protein of interest (POI), Src SH2, PDE6D, NQO2, and Pin1, were amplified from a human placenta cDNA library (TaKaRa BIO) by PCR using appropriate primers. The DNA encoding NLcc and linker sequences, CFRRISGTS-$(GGSGGCGGSGGS)_3$-HM, termed NLcc-linker, was constructed by PCR. The dsDNAs encoding NLcc-linker and POI were ligated into the

MscI/XhoI or MscI/SalI site of pET22b(+). Sequences of the plasmid DNA constructs were verified by dye-terminator sequencing. The plasmid encoding NLmut-SH2 was also constructed in a similar manner. Amino acid sequences were shown:

NLcc-SH2
CFRRISGTSGGSGGCGGSGGSGGSGGCGGSGGSGGSGGCGGSGGSHMDS
IQAEEWYFGKITRRESERLLLNAENPRGTFLVRESETTKGAYCLSVSDF
DNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGLCHR
LTTVCPLEHHHHHH

NLmut-SH2
LFRRISGTSGGSGGCGGSGGSGGSGGCGGSGGSGGSGGCGGSGGSHMDS
IQAEEWYFGKITRRESERLLLNAENPRGTFLVRESETTKGAYCLSVSDF
DNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGLCHR
LTTVCPLEHHHHHH

NLcc-PDE6D
CFRRISGTSGGSGGCGGSGGSGGSGGCGGSGGSGGSGGCGGSGGSHMSA
KDERAREILRGFKLNWMNLRDAETGKILWQGTEDLSVPGVEHEARVPKK
ILKCKAVSRELNFSSTEQMEKFRLEQKVYFKGQCLEEWFFEFGFVIPNS
TNTWQSLIEAAPESQMMPASVLTGNVIIETKFFDDDLLVSTSRVRLFYV
EHHHHHH

NLcc-NQO2
CFRRISGTSGGSGGCGGSGGSGGSGGCGGSGGSGGSGGCGGSGGSHMAG
KKVLIVYAHQEPKSFNGSLKNVAVDELSRQGCTVTVSDLYAMNFEPRAT
DKDITGTLSNPEVFNYGVETHEAYKQRSLASDITDEQKKVREADLVIFQ
FPLYWFSVPAILKGWMDRVLCQGFAFDIPGFYDSGLLQGKLALLSVTTG
GTAEMYTKTGVNGDSRYFLWPLQHGTLHFCGFKVLAPQISFAPEIASEE
ERKGMVAAWSQRLQTIWKEEPIPCTAHWHFGQLEHHHHHH

NLcc-Pin1
CFRRISGTSGGSGGCGGSGGSGGSGGCGGSGGSGGSGGCGGSGGSHMAD

EEKLPPGWEKRMSRSSGRVYYFNHITNASQWERPSGNSSSGGKNGQGEP
ARVRCSHLLVKHSQSRRPSSWRQEKITRTKEEALELINGYIQKIKSGEE
DFESLASQFSDCSSAKARGDLGAFSRGQMQKPFEDASFALRTGEMSGPV
FTDSGIHIILRTE

NLcc-11S-Pin1

CFRRISGGSGGSGGSVFTLEDFVGDWEQTAAYNLDQVLEQGGVSSLLQN
LAVSVTPIQRIVRSGENALKIDIHVIIPYEGLSADQMAQIEEVFKVVYP
VDDHHFKVILPYGTLVIDGVTPNMLNYFGRPYEGIAVFDGKKITVTGTL
WNGNKIIDERLITPDGSMLFRVTINSGGGGTSGGSGGCGGSGGSGGSGG
CGGSGGSGGSGGCGGSGGSHMADEEKLPPGWEKRMSRSSGRVYYFNHIT
NASQWERPSGNSSSGGKNGQGEPARVRCSHLLVKHSQSRRPSSWRQEKI
TRTKEEALELINGYIQKIKSGEEDFESLASQFSDCSSAKARGDLGAFSR
GQMQKPFEDASFALRTGEMSGPVFTDSGIHIILRTE

The sequences corresponding to POI are shown in red character.

NLcc-appended proteins were expressed in BL21(DE3)pG·KJE8 as described above except for the antibiotics (100 μg/mL ampicillin). To obtain the purified proteins (NLcc-SH2, NLcc-PDE6D, and NLcc-NQO2), cells were collected by centrifugation and resuspended in 12 mL of 1 x TES buffer [0.2 M Tris·HCl (pH 8.0), 0.5 mM EDTA, 0.5 M sucrose] at 4°C for 5 min. After that, 18 mL of 1/5 x TES buffer was added and the suspension was incubated at 4°C for 30 min. After centrifugation (10,000 g for 70 min at 4°C), the supernatant was mixed with 4.2 mL of 8 x bind buffer, and loaded onto Ni-NTA agarose resin. Purification of the proteins was performed as described above, and the buffer was exchanged into assay buffer A [100 mM Tris·HCl (pH 7.4), 100 mM NaCl].

To obtain crude preplasmic fraction including NLcc-Pin1 and NLcc-11S-Pin1, cells were collected by centrifugation and periplasmic fraction was extracted using 1 x TES and 1/5 x TES. After centrifugation, the supernatant was collected and stored at -80°C. Concentrations of NLcc-SH2,

NLcc-PDE6D, and NLcc-NQO2 were determined by BCA assay. Concentrations of NLcc-Pin1 and NLcc-11S-Pin1 were estimated from SDS-PAGE analysis by comparison of the band corresponding to NLcc-Pin1 with that corresponding to NLcc-SH2.

IDNCL-ER *in vitro*

For NanoLuc-based IDNCL-ER assay, NLcc-SH2, NLcc-PDE6D, NLcc-NQO2 were diluted with assay buffer A containing 6.25 mM glutathione and 0.125% (v/v) BSA. NLcn-appended ligands were diluted with 0.025% (v/v) Tween 20 aqueous solution. 32 μL of protein solution was mixed with 8 μL of the ligand solution, and then the reaction mixture was incubated at 37°C for 10 min. After that, 5 μL of 50 μM 11S in storage buffer and 5 μL of 100 μM coelenterazine in 10% (v/v) methanol in storage buffer were added. Luminescence intensities were recorded on a Glomax Multi Jr luminometer (Promega).

To assess the interactions of NLcc-Pin1 or NLcc-11S-Pin1 with NLcn-appended ligands, NLcc-Pin1 or NLcc-11S-Pin1 was diluted with assay buffer B [43.75 mM HEPES (pH 7.4), 125 mM NaCl] containing 6.25 mM glutathione and 0.125% (v/v) BSA. NLcc-Pin1 was reacted with NLcn-appended ligands at 37°C for 10 min. After that, 8 μL of 3.125 μM 11S in storage buffer and 2 μL of 250 μM furimazine in methanol were added, and the luminescence intensity was measured. In NLcc-11S-Pin1, 0.5 μL of 250 μM furimazine in methanol were added into 50 μL of the IDNCL reaction solution, and the luminescence intensity was measured.

The dissociation constants were determined by assuming the following equations:

$$\mathrm{L} + \mathrm{P} \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{L{\cdot}P} \xrightarrow{k_L} \mathrm{R}$$
$$+$$
$$\mathrm{G} \xrightarrow{k_G} \mathrm{L'} + \mathrm{P} \qquad \text{---------} \quad (1)$$

$$K_d = \frac{k_{-1}}{k_1} \qquad \text{---------} \quad (2)$$

Where L is NLcn-appended ligand, P is NLcc-appended protein, L·P is the complex, R is the IDNCL product, G is glutathione, L′ is the byproduct generated by the reaction with glutathione, k_1 is the association rate constant, k_2 is the dissociation rate constant, k_L is the rate constant for IDNCL, and k_G is the rate constant for the side reaction by glutathione. Because the complex formation of the ligand with the protein is sufficiently fast compared with the ligation and the side reaction even in the presence of 5 mM glutathione ($k_1 >> k_L$ and k_G), concentration of the complex ([L·P]) is presented as follows:

$$[\mathrm{L{\cdot}P}] = \frac{[\mathrm{P}]_0 + [\mathrm{L}]_0 + K_d - \sqrt{([\mathrm{P}]_0 + [\mathrm{L}]_0 + K_d)^2 - 4[\mathrm{P}]_0[\mathrm{L}]_0}}{2} \qquad \text{---------} \quad (3)$$

where $[\mathrm{P}]_0$ and $[\mathrm{L}]_0$ represent the total concentrations of NLcc-appended protein and NLcn-appended ligand, respectively.

Moreover, when the product concentration is much lower than the complex concentration ([R] << [L·P]), the product concentration is represented as follows:

$$V_0 \times t = [\mathrm{R}] = tk_L[\mathrm{L{\cdot}P}] \qquad \text{----------} \quad (4)$$

Since the product is mixed with excess amount of 11S to measure the luciferase activity, the luminescence intensity proportional to the product concentration.

According to the equations (3) and (4), the luminescence intensity is shown as the following equation:

$$\frac{[\mathrm{L{\cdot}P}]}{[\mathrm{P}]_0} = \frac{L - L_0}{L_{MAX}} \qquad \text{---------} \quad (5)$$

Where L, L_0, L_{MAX} are the luminescence intensities in test solution, in the absence of NLcn-appended ligand, and when NLcc-appended protein was completely comlexed with NLcn-appended ligand, respectively. According to the equations (3) and (5), the following equation (6) was used to calculate the dissociation constants from the luminescence measurements:

$$L = L_0 + L_{MAX} \frac{[P]_0 + [L]_0 + K_d - \sqrt{([P]_0 + [L]_0 + K_d)^2 - 4[P]_0[L]_0}}{2[P]_0} \quad \text{----------} \quad (6)$$

To determine the dissociation constants by competitive binding analysis, NLcn-Dpep1n (800 nM) was reacted with NLcc-Pin1 (10 nM) in the presence of various concentrations of Ac-Dpep1, Ac-Dpep1n, and Ac-Dpep2n. Luciferase activities were measured as described above, and the obtained data were analyzed according to a literature.[3]

Western blotting to confirm IDNCL of NLcn-pYEEI with NLcc-SH2

IDNCL reactions of Bio-NLcn-pYEEI and Bio-NLcn-YEEI with NLcc-SH2 and NLmut-SH2 were carried out as described above. The reaction samples were analyzed by 15% SDS-PAGE, and the proteins were blotted onto a nitrocellulose membrane (GE Healthcare). The membrane was blocked 0.6% (v/v) BSA in TBS-T [20 mm Tris·HCl (pH 7.4), 150 mm NaCl, 0.1% (v/v) Tween 20] on a SNAP i.d. system (Millipore). A horseradish peroxidase-conjugated streptavidin (Proteintech) (1/3000) in TBS-T containing 0.6% BSA was added onto the membrane for 10 min. After washing with TBS-T, the signals were detected by a Supersignal Westfemto HRP substrate (ThermoFisher Scientific).

Assay using cultured cells

HeLa Tet-Off cell line was generated according to the manufacture's protocol by use of pTet-Off Advanced vector (Clontech). HeLa Tet-Off cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) tetracycline reduced fetal bovine serum (FBS), 100 unit/mL penicillin/streptomycin and 100 μg/mL G418 and 2 mM L-glutamine at 37°C in humidified 5% CO_2 atmosphere. The plasmids encoding NLcc-Pin1 was constructed by use of pTRE-tight vector. For the assay using 1 μM NLcn-appended ligands, HeLa Tet-Off cells were seeded in 12-well plates at a density of 8 x 10^4 cells/well and cultured in DMEM supplemented with 10% (v/v) FBS, 2 mM l-glutamine and 100 unit/mL penicillin/streptomycin for 24 h. The expression vector (500 ng/well) was transfected by lipofectamine ltx (2 μL/well). After 4.5 h and 28 h, the medium was exchanged by 1 mL of the medium. After 2 days from the transfection, the cells were washed with PBS (1 mL x 2), and treated with 200 μL of 1 μM NLcn-appended peptide in PBS for 10 min at 37°C. The cells were then washed with PBS (1 mL x 3), and treated with 100 μL of lysis buffer [10 mM HEPES (pH 7.8), 10 mM KCl, 0.1 mM EDTA, 1 mM DTT, 0.1% (v/v) NP-40 substitute] supplemented with a complete protease inhibitor cocktail (Roche) for 30 min at 4°C. The lysates were collected in micro tubes and centrifuged (12,000 g) for 15 min at 4°C. When using cells expressing NLcc-Pin1 and NLcc-PDE6D, 90 μL of the supernatant was mixed with 6 μL of 8.3 μM 11S and 4 μL of 250 μM furimazine. When using cells expressing NLcc-11S-Pin1, 90 μL of the supernatant was mixed with 4 μL of 250 μM furimazine. The luminescence was then recorded. For the assay using several concentrations of the peptides and atorvastatin, the half amounts of the cells and the solutions were used with 24-well plates.

The cell-based IDNCL-ER assay without lysis was performed as follows. HeLa Tet-Off cells were seeded in a 96-well plate at a density of 8 x 10^3

cells/well. After 2 days from the transfection, the cells were washed with PBS (1 mL x 2), and treated with 50 μL of 1 μM NLcn-appended peptide in PBS for 10 min at 37°C. 6 μL of a solution containing trypsin and EDTA (10 x) was added and the plate was incubated at 37°C for 10 min. The solutions with cells were transferred in micro tubes and 1 μL of 500 μM furimazine was added. The luminescence was then recorded.

Western blotting to confirm IDNCL of NLcn-Dpep2 with NLcc-Pin1 in HeLa cells

HeLa Tet-Off cells were seeded in 24-well plates at a density of 4 x 10^4 cells/well and the expression vector was transfected as described above. After 2 days from the transfection, the cells were washed with PBS (0.5 mL x 2), and treated with 100 μL of 250 nM, 25, or 2.5 nM NLcn-Dpep2 in PBS for 10 min at 37°C. The cells were then washed with PBS (1 mL x 3), and treated with 50 μL of lysis buffer. The lysates were collected in micro tubes and centrifuged (12,000 g) for 15 min at 4°C. The supernatant (15 μL) was applied onto 15% SDS-PAGE. As a control, IDNCL reaction of NLcn-Dpep2 (100 or 25 nM) with NLcc-Pin1 (100 nM) *in vitro* was carried out, and the reaction solutions (4 μL) were also conduced. Detection of the proteins bearing biotin moiety was performed as described above.

Stability of phosphopeptides in HeLa cell extracts

HeLa Tet-Off cells were seeded in a 24-well plate at a density of 4 x 10^4 cells/well. After three days, the cells were lysed by 50 μL of lysis buffer in the presence of protease inhibitors described above. After centrifugation, 28 μL of the supernatant was mixed with 2 μL of 300 μM Aoa-Lpep1 or Aoa-Dpep1, and incubated at 37°C. 10 μL of the reaction solution was injected onto a Daisopack

SP300 C4-BIO (4.6 x 150 mm) with a linear gradient of acetonitrile/0.1% aqueous TFA (20-50% over 30 min) at 40°C.

References

1. J. B. Blanco-Canosa and P. E. Dawson, *Angew. Chem. Int. Ed.*, 2008, **47**, 6851-6855.
2. M. Kawasse, M. Fujioka and T. Takahashi, *ChemBioChem*, 2021, **22**, 577-584.
3. M. H. A. Roehrl, J. Y. Wang, G. Wagner, *Biochemistry*, 2004, **43**, 16056-16066.

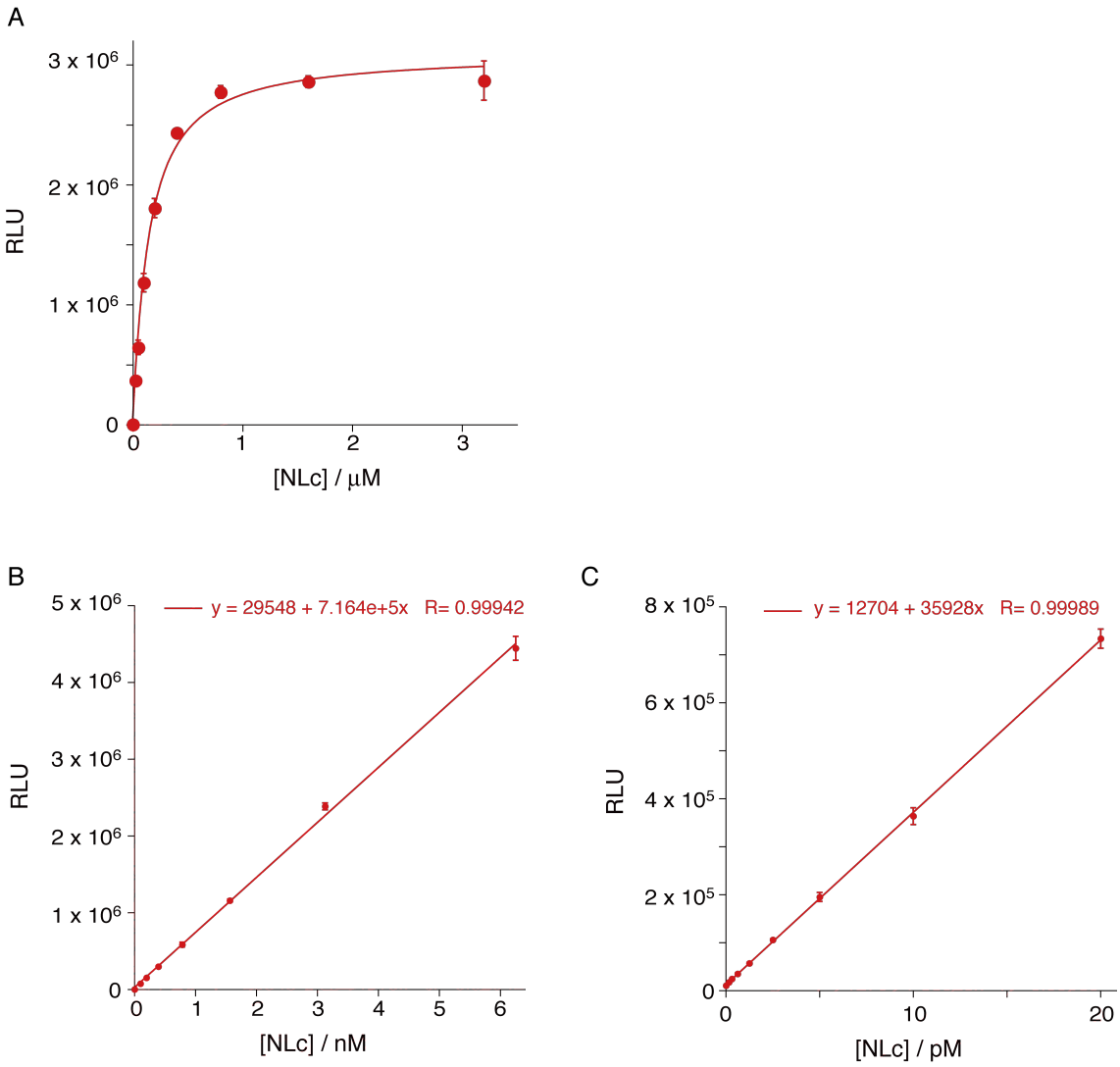


Fig. S1 Complementation between 11S and NLc peptide. (A) Luminometric titration curve of 11S of the luminescence intensity versus the NLc peptide concentration. Various concentrations of NLc was mixed with 11S (50 nM) in 80 mM Tris·HCl (pH 7.4) containigng 80 mM NaCl, 0.1% (v/v) BSA, 5 mM glutathione, and 0.001% (v/v) Tween 20. Coelenterazine (10 μM) was used. (B) and (C) Luminescence intensity as a function of NLc concentration. Various concentrations of NLc was mixed with (B) 11S (5 μM) and coelenterazine (10 μM) and (C) 11S (0.5 μM) and furimazine (10 μM) in the buffer shown as in (A).

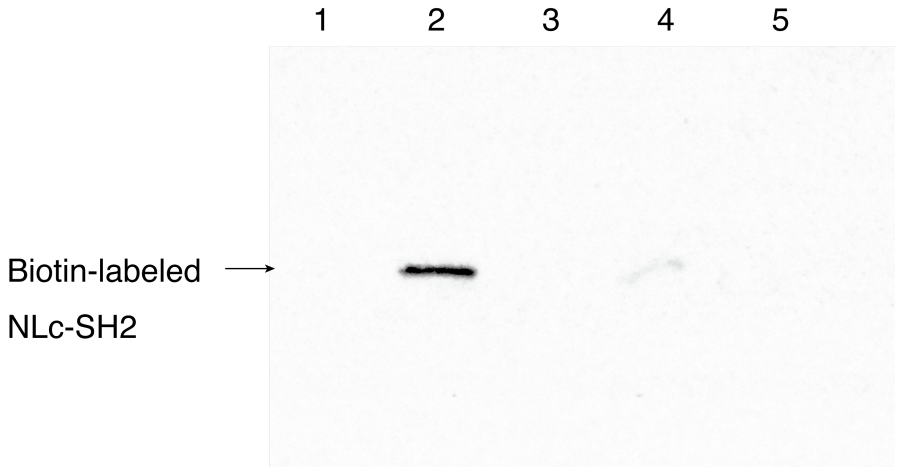


Fig. S2 Western blotting analysis of IDNCL. Biotin-appended NLcn peptides (1.6 μM) was reacted with NLcc-SH2 or NLmut-SH2 (100 nM) in the buffer shown as Fig. S1 (A) for 30 min at 37°C. lane 1, NLcc-SH2 alone; lane 2, reaction of NLcn-pYEEI with NLcc-SH2; lane 3, reaction of NLcn-YEEI with NLcc-SH2; lane 4, reaction of NLcn-pYEEI with NLcc-SH2 in the presence of 100 μM Ac-pYEEI-NH_2; lane 5, reaction of NLcn-pYEEI with NLmut-SH2.

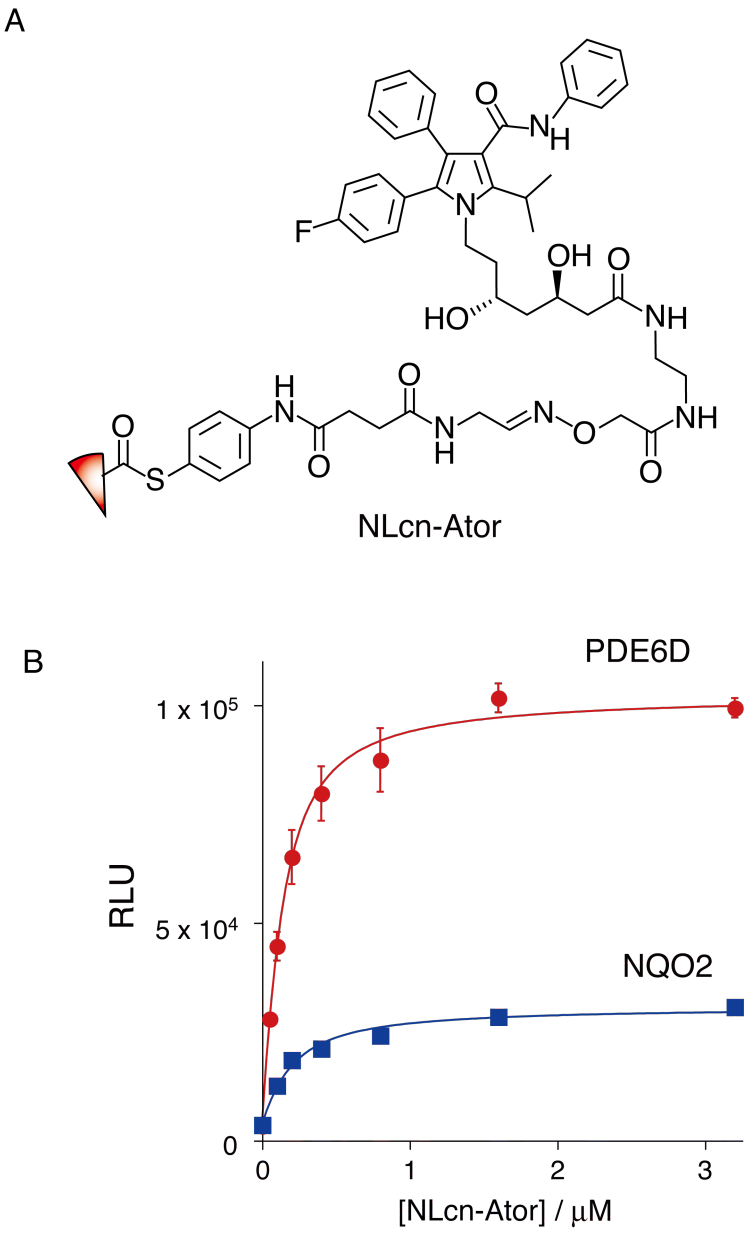


Fig. S3 (A) Schematic structure of NLcn-appended atorvastatin, NLcn-Ator. (B) Luminescence changes as a function of NLcn-Ator concentration in the reaction of NLcn-Ator with NLcc-PDE6D (100 nM, red filled circle) and NLcc-NQO2 (100 nM, blue filled square).

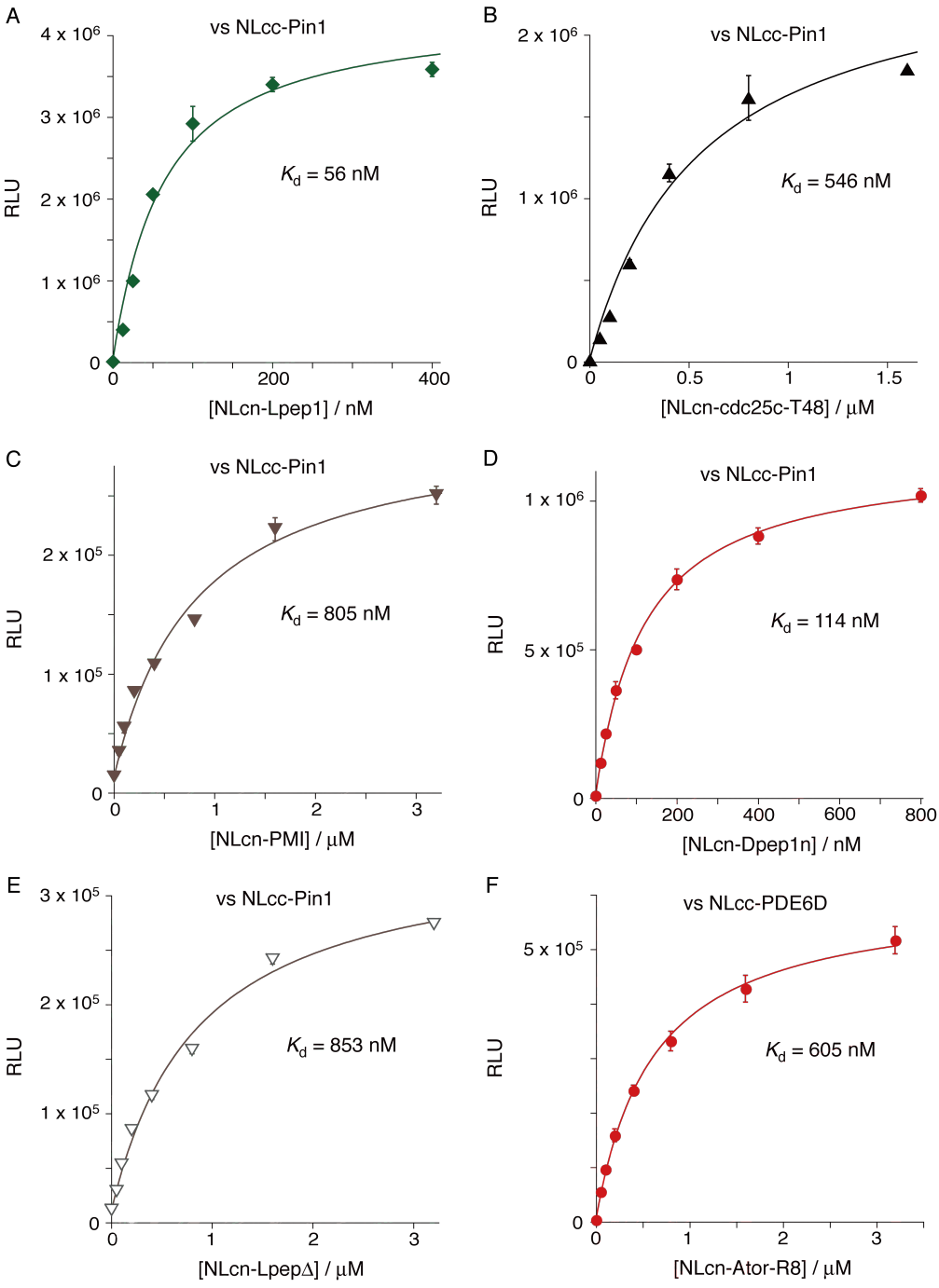

A
vs NLcc-Pin1
RLU
4 x 10^6
3 x 10^6
2 x 10^6
1 x 10^6
0
K_d = 56 nM
0
200
400
[NLcn-Lpep1] / nM
B
vs NLcc-Pin1
RLU
2 x 10^6
1 x 10^6
0
K_d = 546 nM
0
0.5
1
1.5
[NLcn-cdc25c-T48] / µM
C
vs NLcc-Pin1
RLU
2 x 10^5
1 x 10^5
0
K_d = 805 nM
0
1
2
3
[NLcn-PMI] / µM
D
vs NLcc-Pin1
RLU
1 x 10^6
5 x 10^5
0
K_d = 114 nM
0
200
400
600
800
[NLcn-Dpep1n] / nM
E
vs NLcc-Pin1
RLU
3 x 10^5
2 x 10^5
1 x 10^5
0
K_d = 853 nM
0
1
2
3
[NLcn-LpepΔ] / µM
F
vs NLcc-PDE6D
RLU
5 x 10^5
0
K_d = 605 nM
0
1
2
3
[NLcn-Ator-R8] / µM

Fig. S4 *In vitro* IDNCL-ER assay to detect the interactions of NLcc-Pin1 with (A) NLcn-Lpep1, (B) NLcn-cdc25c-T48, (C) NLcn-PMI, (D) NLcn-Dpep1n, and (E) NLcn-LpepD, and (F) NLcc-PDE6D with NLcn-Ator-R8. [NLcc-Pin1] = 10 nM, [NLcc-PDE6D] = 10 nM.

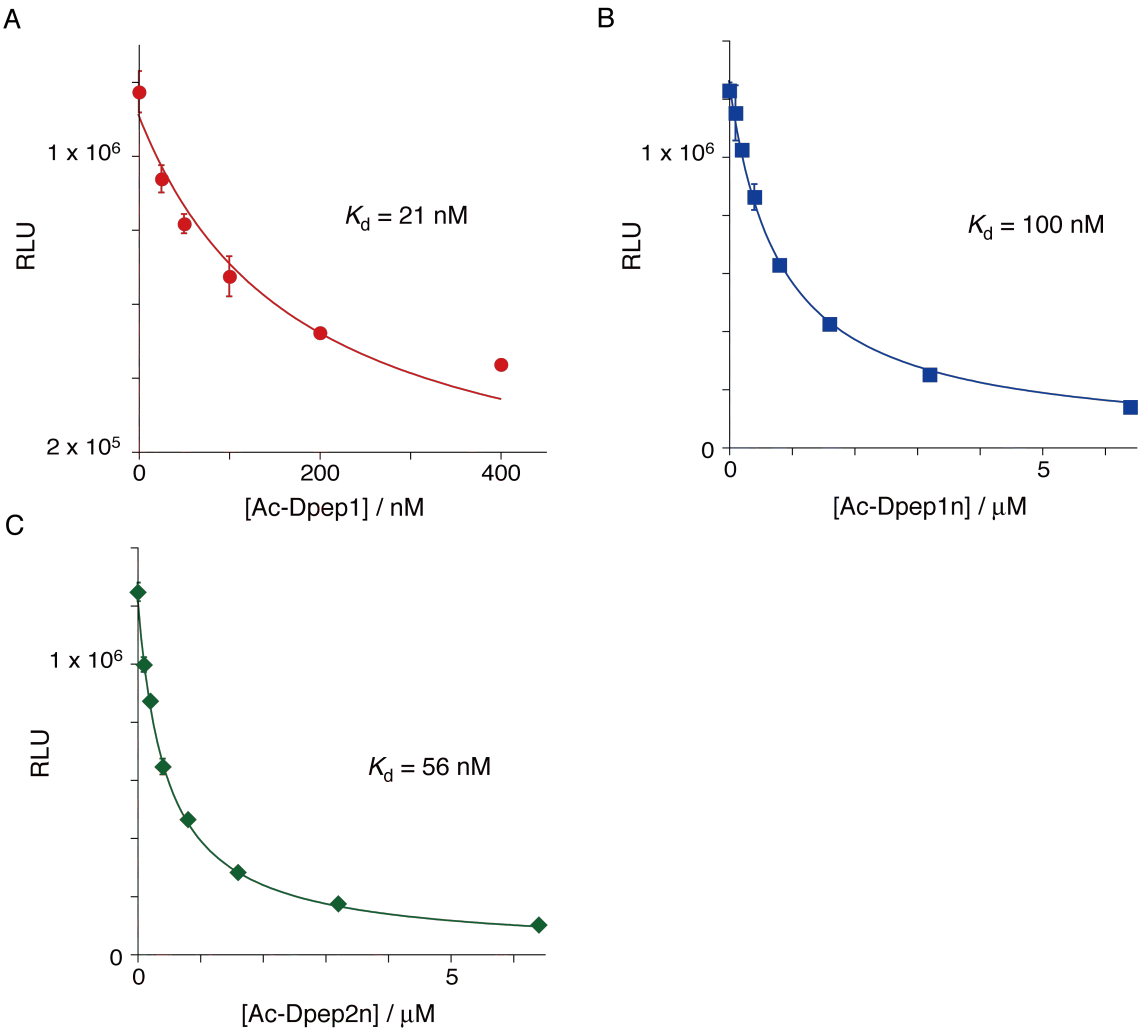


Fig. S5 Competitive binding assay using IDNCL-ER for the interaction between NLcn-Dpep1n and NLcc-Pin1 in the presence of various concentrations of (A) NLcn-Dpep1, (B) NLcn-Dpep1n, and (C) NLcn-Dpep2n. [NLcn-Dpep1n] = 800 nM, [NLcc-Pin1] = 10 nM.

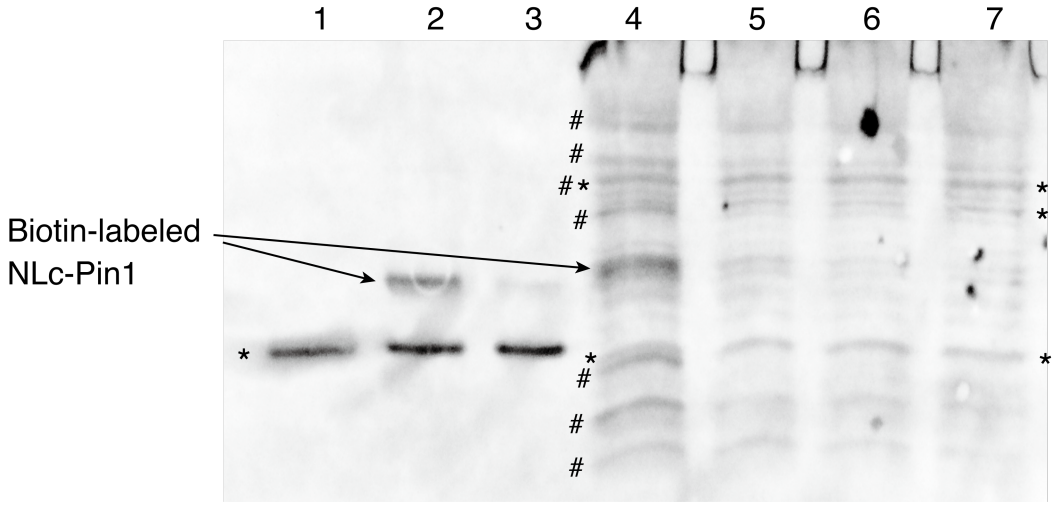


Fig. S6 Western blotting analyses of IDNCL *in vitro* (lane 1 – 3) and inside cells (lane 4 – 7). lane 1, NLcc-Pin1 alone; lane 2, reaction of NLcn-Dpep2 (100 nM) with NLcc-Pin1 (100 nM); lane 3, reaction of NLcn-Dpep2 (25 nM) with NLcc-Pin1 (100 nM); lane 4, HeLa Tet-Off cells expressing NLcc-Pin1 treated with NLcn-Dpep2 (250 nM); lane 5, cells treated with NLcn-Dpep2 (25 nM); lane 6, cells treated with NLcn-Dpep2 (2.5 nM); lane 7, cells treated without the peptide. Asterisk indicates the bands corresponding to non-specific interactions with horseradish peroxidase conjugated with streptavidin. Octothorpe indicates the bands corresponding to off-target proteins that react with NLcn-Dpep2 non-specifically.

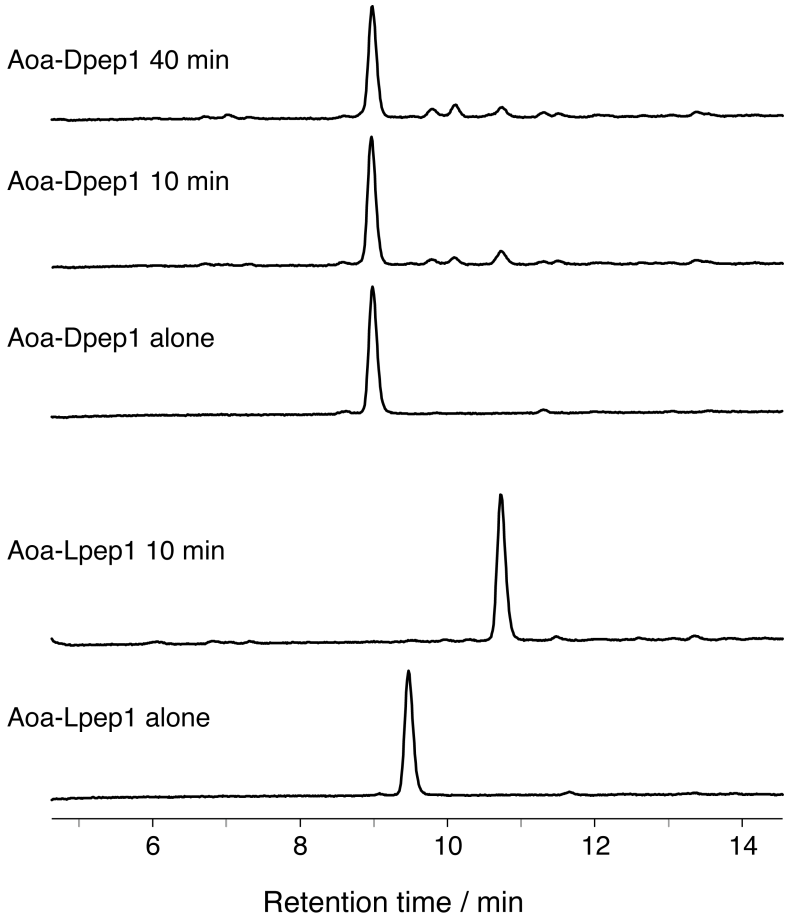


Fig. S7 HPLC analyses of the stability of phosphopeptides. Aoa-Lpep1 and Aoa-Dpep1 were mixed with HeLa cell extracts and incubated at 37°C.

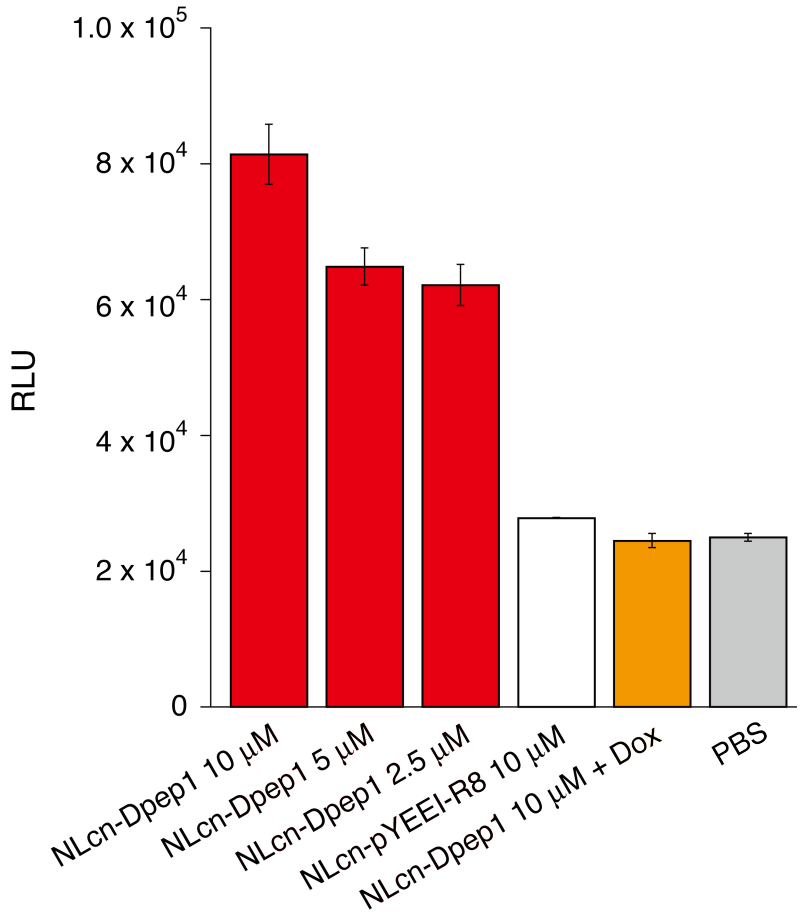


Fig. S8 IDNCL-ER assay of NLcn-appended ligands with NLcc-Pin1 using 12-well plates. NLcn-Dpep1 (10, 5, and 2.5 μM) and NLcn-pYEEI-R8 (10 μM) in the absence and in the presence of doxycycline (20 ng/mL) were incubated with NLcc-Pin1-expressing cells for 10 min at 37°C. The lysates from the cells were mixed with 11S (0.5 μM) and furimazine (10 μM) and the luminescence was measured. Error bars represent the mean ± standard deviation of three samples.

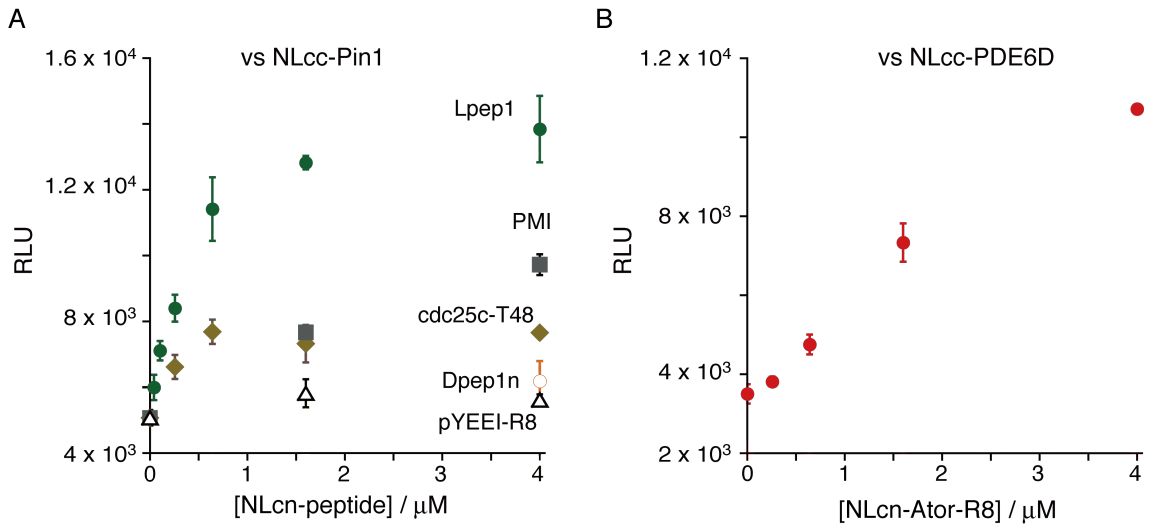


Fig. S9 IDNCL-ER assay of NLcn-appended ligands with (A) NLcc-Pin1 and (B) NLcc-PDE6D. Concentration-dependent responses of NLcn-appended peptides and NLcn-Ator-R8 in IDNCL-ER assay with NLcc-Pin1- and NLcc-PDE6D-expressing cells

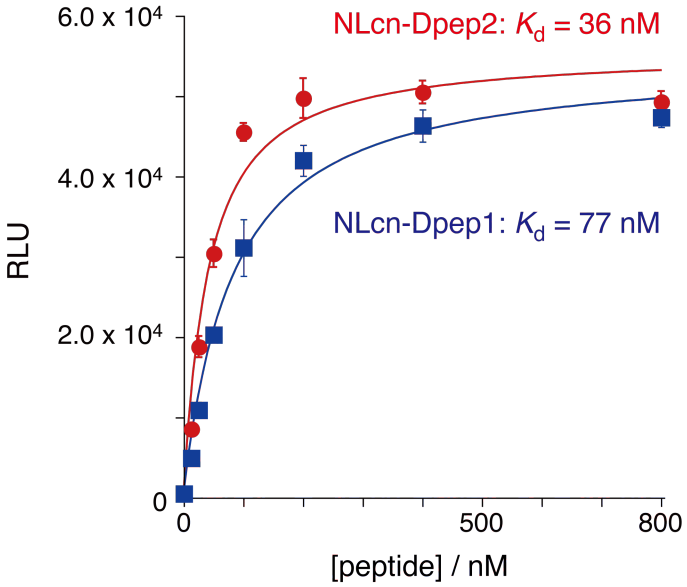


Fig. S10 *In vitro* IDNCL-ER assay to detect the interactions of NLcc-11S-Pin1 with NLcn-Dpep1 (blue filled square) and NLcn-Dpep2 (red filled circle).

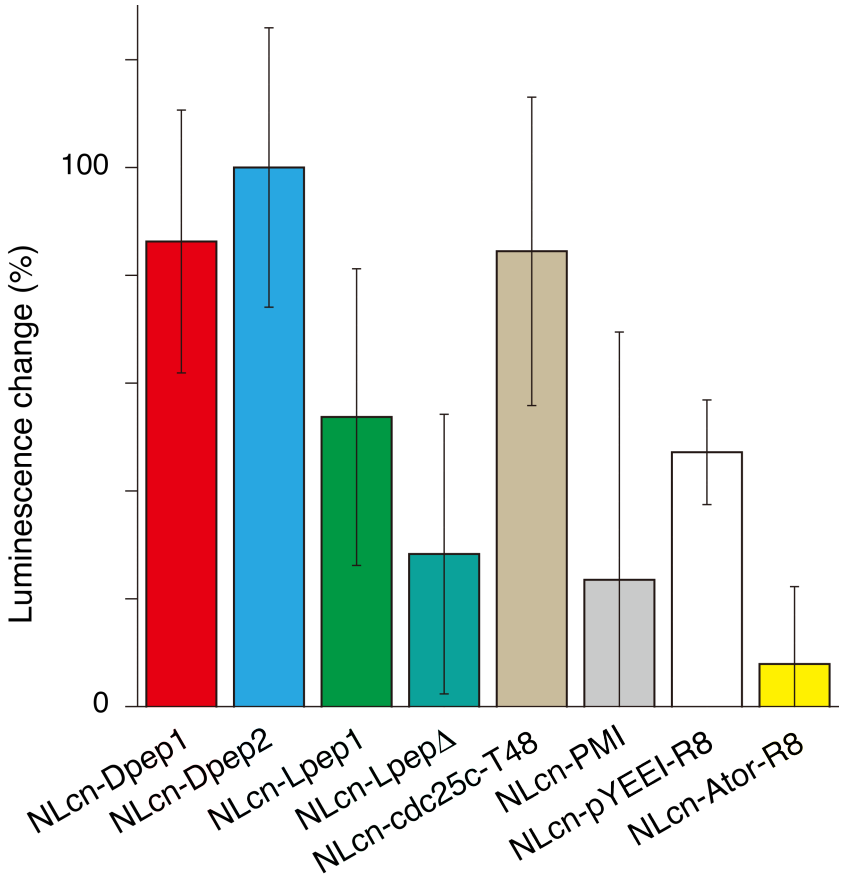


Fig. S11 IDNCL-ER assay of NLcn-appended ligands with NLcc-11S-Pin1 using 96-well plates. HeLa Tet-Off cells expressing NLcc-11S-Pin1 were treated with 1 μM NLcn-appended ligands at 37°C for 10 min. Cells were detached from the cells and luminescence was measured in the presence of 10 μM furimazine.