Supplementary Information

Crystallographic analysis of G-clamp-RNA complex assisted by large scale RNA-binding profile

Ryosuke Nagasawa,[ab] Kazumitsu Onizuka,*[abc] Karen Kawamura,[d] Kosuke Tsuzuki,[ab] Hirotaka Murase,[ab] Kaoru R. Komatsu,[e] Emi Miyashita,[e] Hirohide Saito,[e] Jiro Kondo,*[d] and Fumi Nagatsugi*[ab]

[a] Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577, Japan.

[b] Department of Chemistry, Graduate School of Science, Tohoku University, Aoba-ku, Sendai 980-8578, Japan.

[c] Division for the Establishment of Frontier Sciences of Organization for Advanced Studies, Tohoku University, Aoba-ku, Sendai, Miyagi 980-8577, Japan.

[d] Department of Materials and Life Sciences, Faculty of Science and Technology, Sophia University, Tokyo 102-8554, Japan.

[e] Center for iPS Cell Research and Application (CiRA), Kyoto University, 53 Kawahara-cho, Shogoin, Sakyoku, Kyoto, 606-8507, Japan.

Table of Contents

Experimental procedure

Material and methods

General chemicals were purchased from FUJIFILM Wako Pure Chemical, the Tokyo Chemical Industry, Kanto Chemical or Aldrich. Target RNAs were purchased from JBioS (Japan). ^{1}H NMR spectra (600 MHz) were recorded using Bruker AVANCE III 600 spectrometer. ^{13}C NMR spectra (151 MHz) were recorded using Bruker AVANCE III 600 spectrometer. High-resolution electrospray mass analysis was performed using a Bruker MicrOTOF-Q II. HPLC purification was performed with a JASCO HPLC System (CO-631A) using a reverse-phase C_{18} column (COSMOSIL 5C_{18}-AR-II, Nacalai Tesque, 10×250 mm for ligand purification).

RNA secondary structure prediction and visualization

The secondary structures of single-stranded and double-stranded RNA were predicted by RNAfold and RNAcofold v. 2.5.1 in the ViennaRNA package[1], respectively, with the temperature set to 25 °C. The forna website[2] was used to generate illustrations of the secondary structures of single-stranded RNA predicted.

Fluorescence binding assay

A solution (100 μL) of the binder (0.1 μM for G-clamp-N_3, 0.01 or 0.1 μM for G-clamp-NH_2) in 1x Binding buffer (5% DMSO, 20 mM phosphate, 20 mM NaCl and 80 mM KCl) was transferred to a micro quartz cell with a 1-cm path length. Serial aliquots of a concentrated solution of RNA in 1x buffer was added to the binder solution and allowed to equilibrate for 1 min. The excitation wavelength was set at 360 nm for both binders, and the emission was recorded at 20 °C. Fluorescence measurements were performed with a JASCO-6500 spectrofluorometer (JASCO, Tokyo, Japan).

The data from the titrations were analyzed according to the independent-site model by non-linear fitting to Equations (1), in which F_0 is the initial fluorescence intensity in the absence of RNA, Q (= F_{max}/F_0) is the fluorescence enhancement upon saturation, A = K_D/C_{ligand} and X = nC_{RNA}/C_{ligand} (n is the putative number of binding sites on RNA and n = 2 was used for pre-mir-125a-X-ray, n =1 was used for the other RNAs)[3]. The parameters Q and X were determined by KaleidaGraph (Synergy Software, PA). The K_D values in the main text show the mean values of three independent experiments.

$$F/F_0 = 1+(Q-1)/2\{A+1+X-[(X+1+A)^2-4X]^{1/2}\} \quad (1)$$

SPR analysis

Immobilization: 5′-biotinylated RNA (pre-mir-125a) was diluted to 1 μM in 1× Binding buffer (20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl), and the solution was heated at 95 °C for 5 min and cooled on ice. The folded RNAs were injected over a streptavidin-coated sensor chip (Series S Sensor chip SA, Cytiva) at 60 μL/min to reach an immobilized level of 1344 RU.

Binding analysis by single-cycle kinetics: the RNA binder (G-clamp-N_3 or G-clamp-NH_2) in 1× Binding buffer (1% DMSO, 20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl) was injected at increasing concentrations (200, 400, 600, 800, and 1000 nM for G-clamp-N_3, 10, 20, 30, 40, and 50 nM for G-clamp-NH_2) to the RNA-immobilized sensor surface without a regeneration step between each concentration. Each of the RNA binder was injected with a flow rate of 60 μL/min, contact time of 30 s, and dissociation time of 120 s using the running buffer at 25 °C. All

sensorgrams were corrected by subtracting the blank flow cell and the buffer injection responses. All kinetics were obtained by Biacore T200 evaluation software. Three technical replicates were conducted.

Crystallization, X-ray data collection, structure determination, and structure refinement

The RNA fragment pre-mir125a-X-ray (XR hereafter) was designed to fold as a pseudo-self-complementary duplex containing two internal loops of pre-mir125a. The 5'-terminal UU residues are added to facilitate crystal packing. The RNA XR was synthesized using a DNA/RNA synthesizer NTS M-2-MX (Nihon Techno Service). After deprotection, the RNA XR was purified using denaturing 20% polyacrylamide gel electrophoresis and subsequently desalted by reversed-phase chromatography using the Sep-Pak C18 Classic Cartridge (Waters).

Crystallizations were performed both in the presence and absence of G-clamp-NH_2. Prior to co-crystallization, 2 mM RNA XR solution was mixed with the same volume of 4 mM G-clamp-NH_2 solution. For the crystallization of RNA in the absence of G-clamp-NH_2, 1 mM RNA XR solution was used. Crystallizations were performed at 293K by the sitting-drop vapor diffusion method by mixing 0.2 µl of XR/G-clamp-NH_2 or XR solution and 0.2 µl of crystallization solution of our homemade crystallization screening kits, which was equilibrated against 30 µl of a reservoir solution. A single crystal of the XR/G-clamp-NH_2 complex was obtained in a condition containing 50 mM HEPES (pH 7.0), 10 mM spermine tetrahydrochloride, 200 mM potassium chloride, 25 mM magnesium sulfate and 20% (v/v) polyethylene glycol 200. The crystal was soaked in a cryoprotectant containing 40% (v/v) 2-methyl-2,4-pentandiol and 10 mM hexammine cobalt chloride for several 10 seconds and then flash frozen in liquid nitrogen. Crystals of XR without G-clamp-NH_2 were obtained in several conditions, but none of them showed diffraction with high resolution sufficient for structural analysis.

An X-ray dataset was collected at 100K with synchrotron radiation at the BL-17A beamlines of the Photon Factory (Tsukuba, Japan). The dataset was processed using the XDS program.[4] The statistics of data collection and the crystal data are summarized in Table S2. The phase determination of the XR/G-clamp-NH_2 complex was attempted with the molecular replacement program *Phaser* from the *Phenix* suite [5, 6] using several crystal structures of the ribosomal A site as models. The molecular structure of the XR/G-clamp-NH_2 complex was constructed and manipulated with the program *Coot*.[7, 8] The atomic parameters were refined using the *phenix.refine* program of the *Phenix* suite through a combination of simulated-annealing, crystallographic conjugate gradient minimization refinements, and B-factor refinements.[5, 9] The statistics of structure refinement are summarized in Table S2. Molecular drawings were made using PyMOL.[10] The atomic coordinate and experimental data of the XR/G-clamp-NH_2 complex have been deposited in the Protein Data Bank (PDB) with the ID code 9J6P.

Footprinting assay using RNase T1

10 µM of 5'FAM labeled pre-mir-125a was folded in 1x Binding buffer (20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl) by heating at 95℃ for 5 min and rapidly cooling on ice. To a solution of 1 µL of the folded RNA, 2.5 µL of H_2O, 4.5 µL of 2x Binding buffer and compound (0-10 µM in 10% DMSO) were added and incubated at 25℃ for 60 min. 1 µL of RNase T1 (Worthington) was added to a final concentration of 108 ng/mL and samples were incubated at 25℃ for an additional 5 min. The reactions were stopped by adding 10 µL of Gel Loading Buffer (80% (v/v) formamide, 10 mM EDTA). To generate the ladder, 5'FAM labeled pre-mir-125a was hydrolyzed under alkaline condition (25 mM carbonate solution (pH 10), 1 mM EDTA) for 12 min at 95℃. Cleavage products were resolved on a denaturing gel (16% (w/v) polyacrylamide, 20% (v/v) formamide, 7.5 M urea), which was quantified by Image

Lab (Bio-Rad). Three independent experiments were conducted. Two-tailed Student's T test was made by Scipy v. 1.7.3 in python 3.7.

Melting temperature (T_m) measurements

T_m curves of pre-mir-125a (1 μM) with G-clamp-N_3 or G-clamp-NH_2 (0, 1, 2, and 4 μM) in 1x Binding buffer (1% DMSO, 20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl) were recorded on a JASCO V-730 spectrophotometer (JASCO, Tokyo, Japan) equipped with a PAC-743 temperature controller and a 1 cm path-length cell. The absorbance of the samples was monitored at 260 nm from 20℃ to 95℃ with a heating rate of 1 ℃ /min. T_m values were calculated by the second derivative method. Three technical replicates were conducted.

Table S1. RNA sequences in this study.

Name	Sequence
Pre-mir-125a (WT)	AGCCUGUGAGGACAUCCAGGGUCACAGGCU
Pre-mir-125a-X-ray (XR)	UUGCGGGGUCCCGGGAGGACCGC
Pre-mir-125a-Δinternal (Δinternal)	AGCCUGUGAAUCCAUCACAGGCU
5'FAM-labeled pre-mir-125a	FAM-AGCCUGUGAGGACAUCCAGGGUCACAGGCU
5'-biotinylated pre-mir-125a	biotin-AGCCUGUGAGGACAUCCAGGGUCACAGGCU

Table S2. Crystal data and statistics of data collection and structure refinement

Crystal code	XR/G-clamp-NH_2 complex
Crystal data	
Space group	*R*32
Unit cell (Å, °)	*a* = *b* = 46.2, *c* = 336.2
No. of dsRNA in AU [a]	1
Data collection	
Beamline	BL17A (PF)
Wavelength (Å)	1.605
Resolution (Å)	39.7-2.5
of the outer shell (Å)	2.6-2.5
Unique reflections	9098
Completeness (%)	97.2
in the outer shell (%)	94.2
R_{merge} [b] (%)	-
in the outer shell (%)	-
R_{anom} [b] (%)	9.0
in the outer shell (%)	34.6
Redundancy	10.3
in the outer shell	10.3
$I/\sigma(I)$	19.1
in the outer shell	8.2
Structure refinement	
Resolution range (Å)	39.7-2.5
Used reflections	9010
R-factor [c] (%)	20.0
R_{free} [d] (%)	26.8
R.m.s.d. bond length (Å)	0.007
R.m.s.d. bond angles (°)	1.2

[a] Number of double-stranded RNA in the asymmetric unit.

[b] $R_{merge} = 100 \times \Sigma_{hklj}|I_{hklj} - \langle I_{hklj}\rangle| / \Sigma_{hklj}\langle I_{hklj}\rangle$.
$R_{anom} = 100 \times \Sigma_{hklj}|I_{hklj}(+) - I_{hklj}(-)| / \Sigma_{hklj}[I_{hklj}(+) + I_{hklj}(-)]$.

[c] *R*-factor $= 100 \times \Sigma||F_o| - |F_c|| / \Sigma|F_o|$, where $|F_o|$ and $|F_c|$ are optimally scaled observed and calculated structure factor amplitudes, respectively.

[d] Calculated using a random set containing 10% of observations.

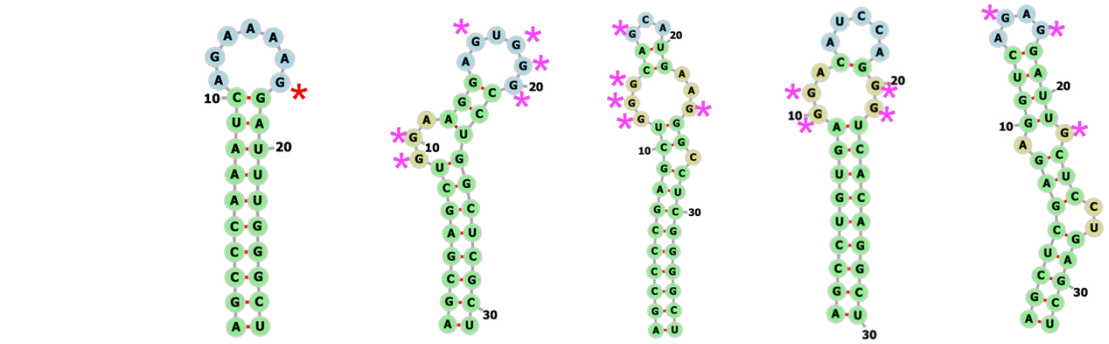


Figure S1. Predicted secondary RNA structures with putative binding sites of G-clamp derivatives. The secondary RNA structures were calculated by RNAfold at 25℃. Magenta stars and red star on G in single-stranded region indicate putative binding sites and identified binding site of G-clamp derivatives, respectively. The rank was obtained by FOREST previously.[11]

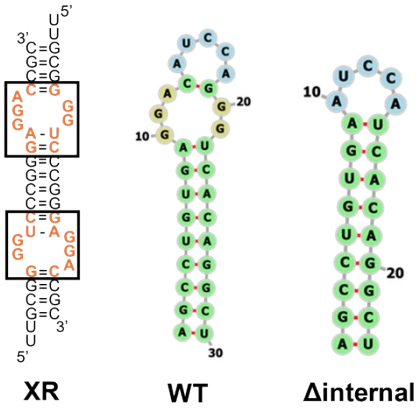


Figure S2. Secondary structure of RNA in this study. Secondary structure of pre-mir-125a-X-ray (XR) is predicted by RNAcofold (-T 25) and the other secondary structures are predicted by RNAfold (-T 25).

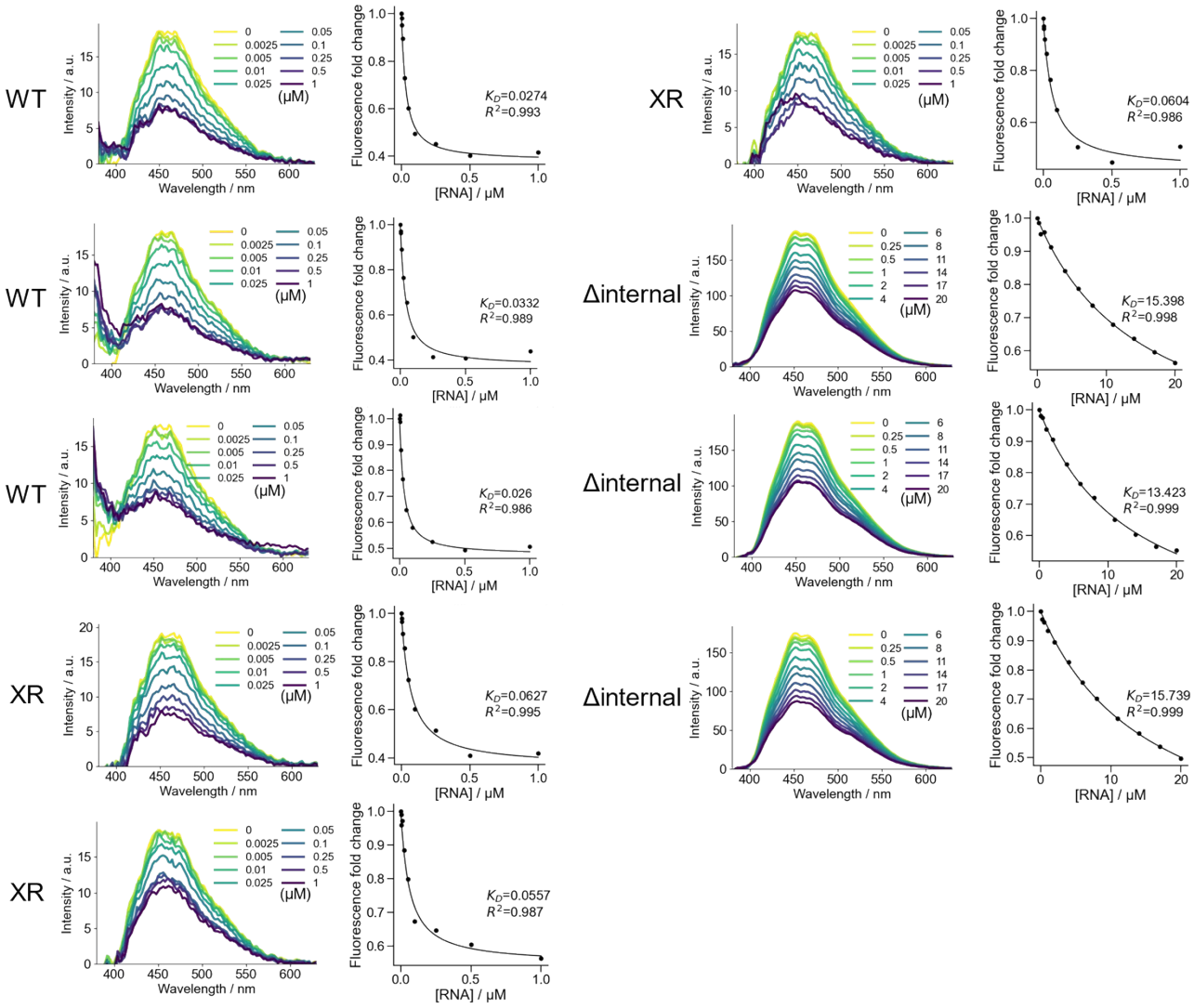


Figure S3. Fluorescence titrations to determine the dissociation constants (K_D) of G-clamp-NH_2 and pre-mir-125a derivatives. Fluorescence titration spectra were measured using G-clamp-NH_2 (0.01 for WT and XR or 0.1 μM for Δinternal) upon addition of RNA (0-1, 20 μM) in 1×Binding buffer (5% DMSO, 20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl). Conditions ; λex : 360 nm, λem : 452 nm, temperature : 20 ℃. The data and K_D values obtained by three independent experiments are shown.

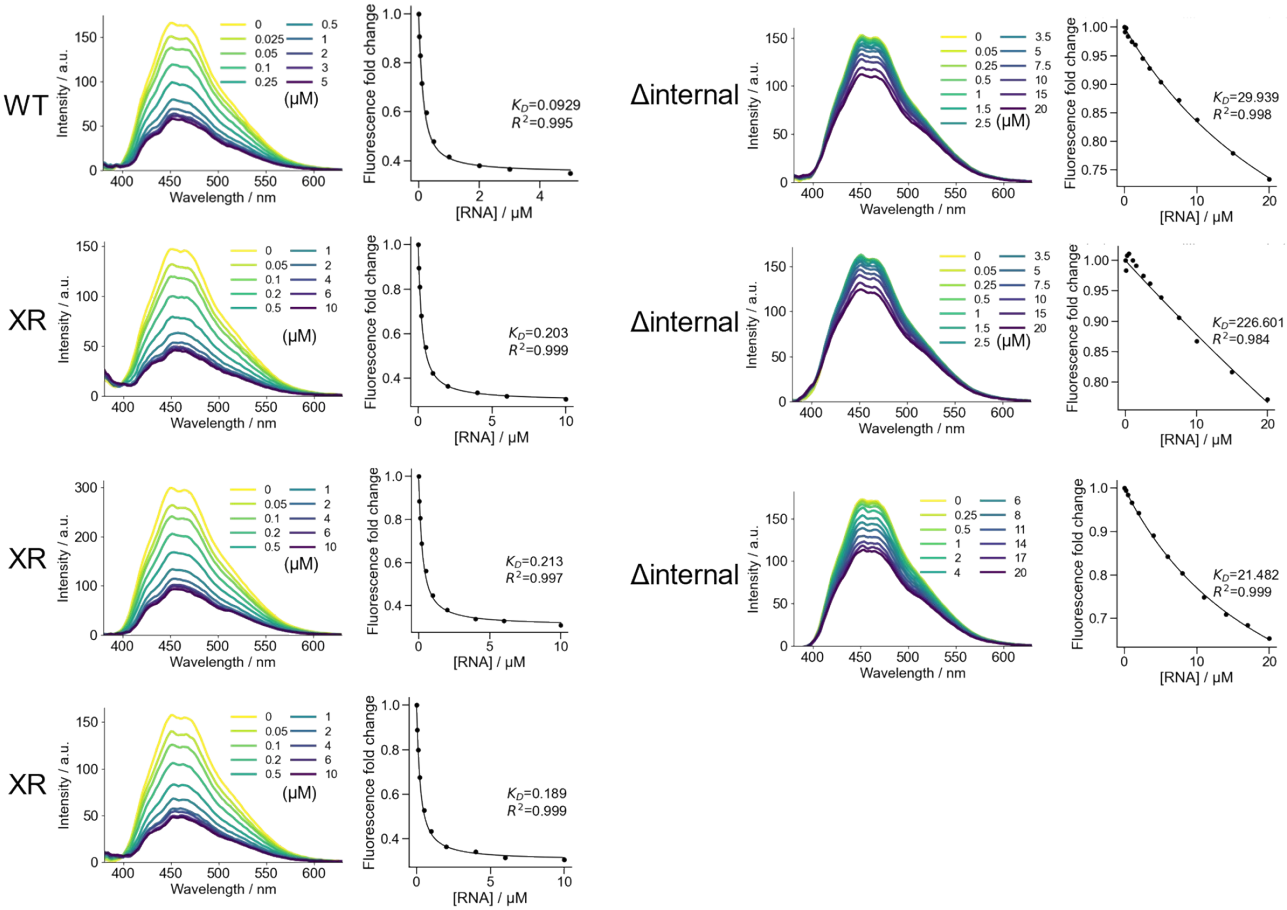


Figure S4. Fluorescence titrations to determine the dissociation constants (K_D) of G-clamp-N_3 and pre-mir-125a derivatives. Fluorescence titration spectra were measured using G-clamp-N_3 (0.1 μM) upon addition of RNA (0-5,10 or 20 μM) in 1×Binding buffer (5% DMSO, 20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl). Conditions ; λex : 360 nm, λem : 452 nm, temperature : 20 ℃. The data and K_D values obtained by three independent experiments are shown. For WT, other two experiments were reported previously.[11]

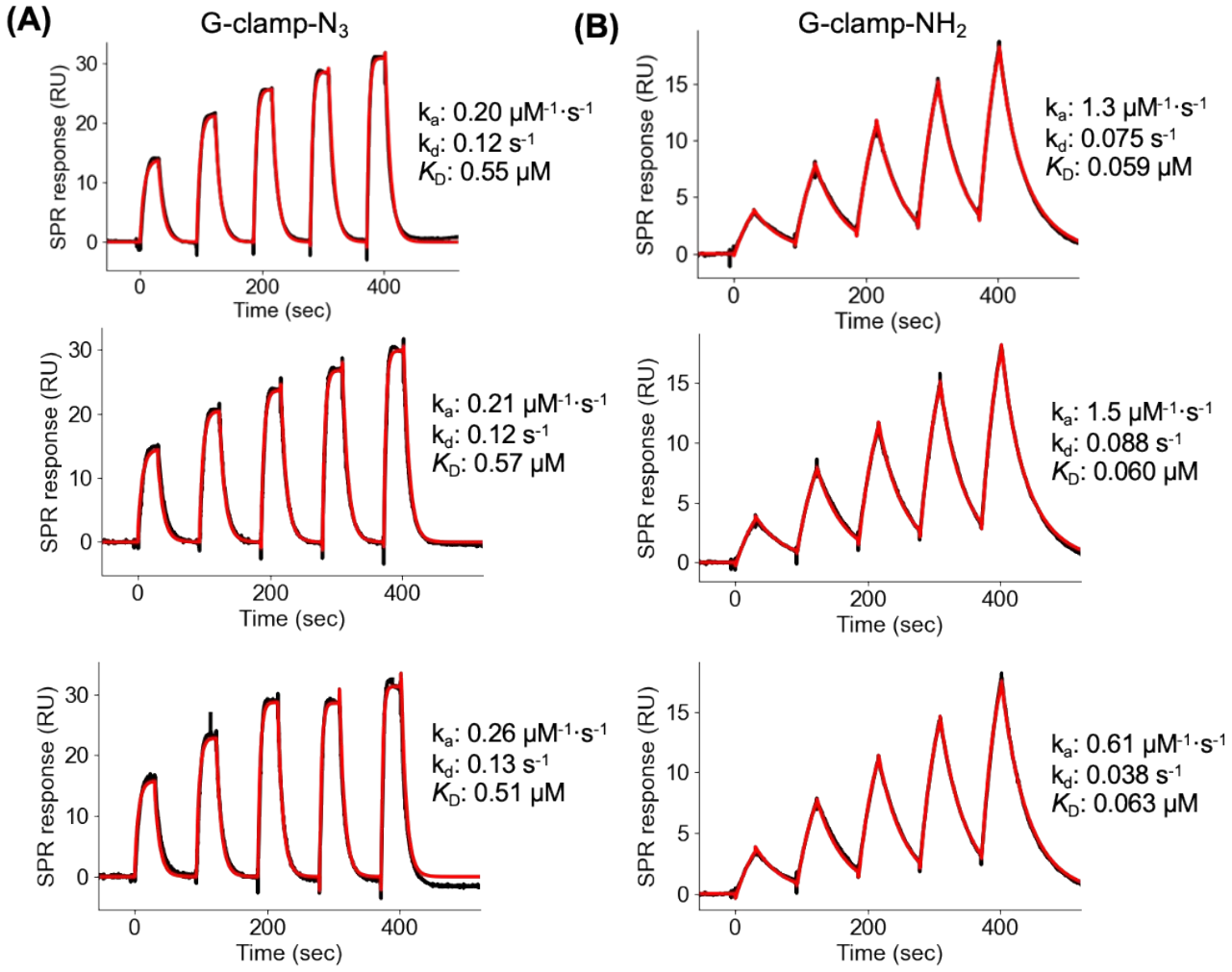


Figure S5. Surface plasmon resonance (SPR) analysis. (**A**, **B**) The SPR sensorgrams of (**A**) G-clamp-N_3 (200, 400, 600, 800, and 1000 nM) and (**B**) G-clamp-NH_2 (10, 20, 30, 40, and 50 nM) with 5'-biotinylated pre-mir-125a. The black line represents the SPR response with the RNA binder and the red line represents the fitting curve. The binding kinetics were determined in single-cycle kinetic mode. Three technical replicates were conducted.

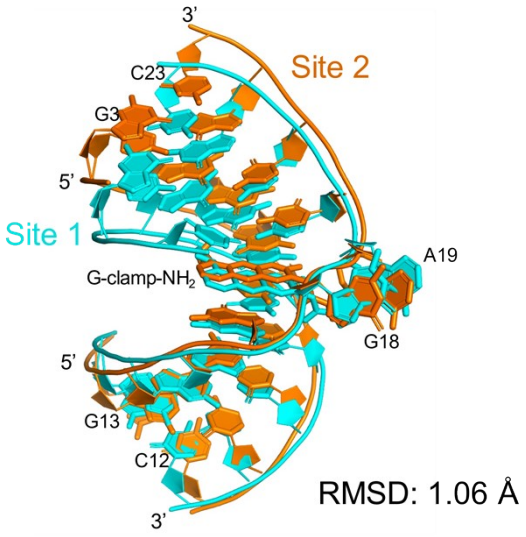


Figure S6. Comparison of site 1 and 2 of the XR-G-clamp-NH_2 complex. The half of the complex including sites 1 in cyan are superposed the half of the complex including sites 2 in orange. The superposition was conducted by align command in PyMOL.

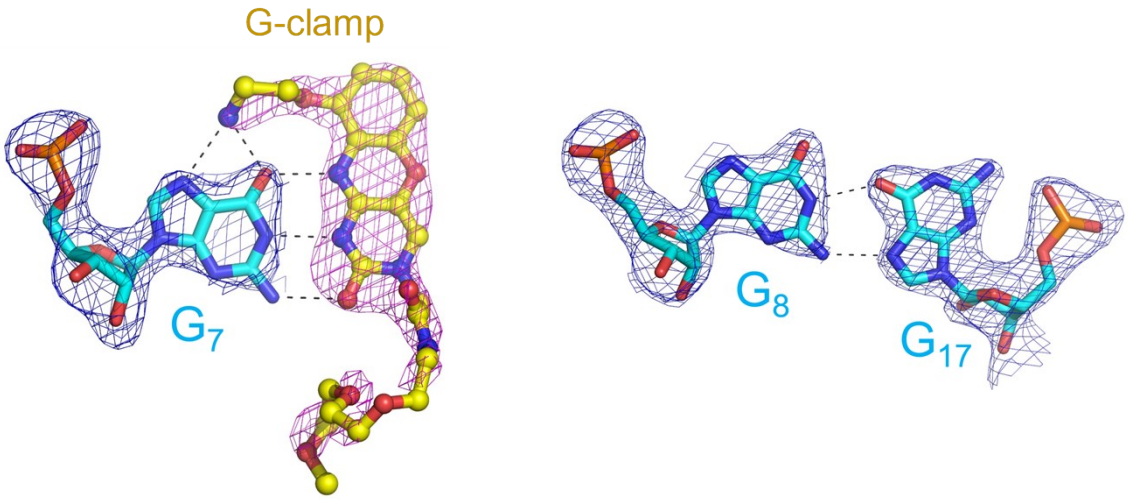


Figure S7. Local $2|F_o| - |F_c|$ (blue: 1 σ contour level) and omit (magenta: 3 σ contour level) maps for the pseudo base pair between G7 and G-clamp and for the *cis* Watson-Crick/Hoogsteen G8-G17 base pair.

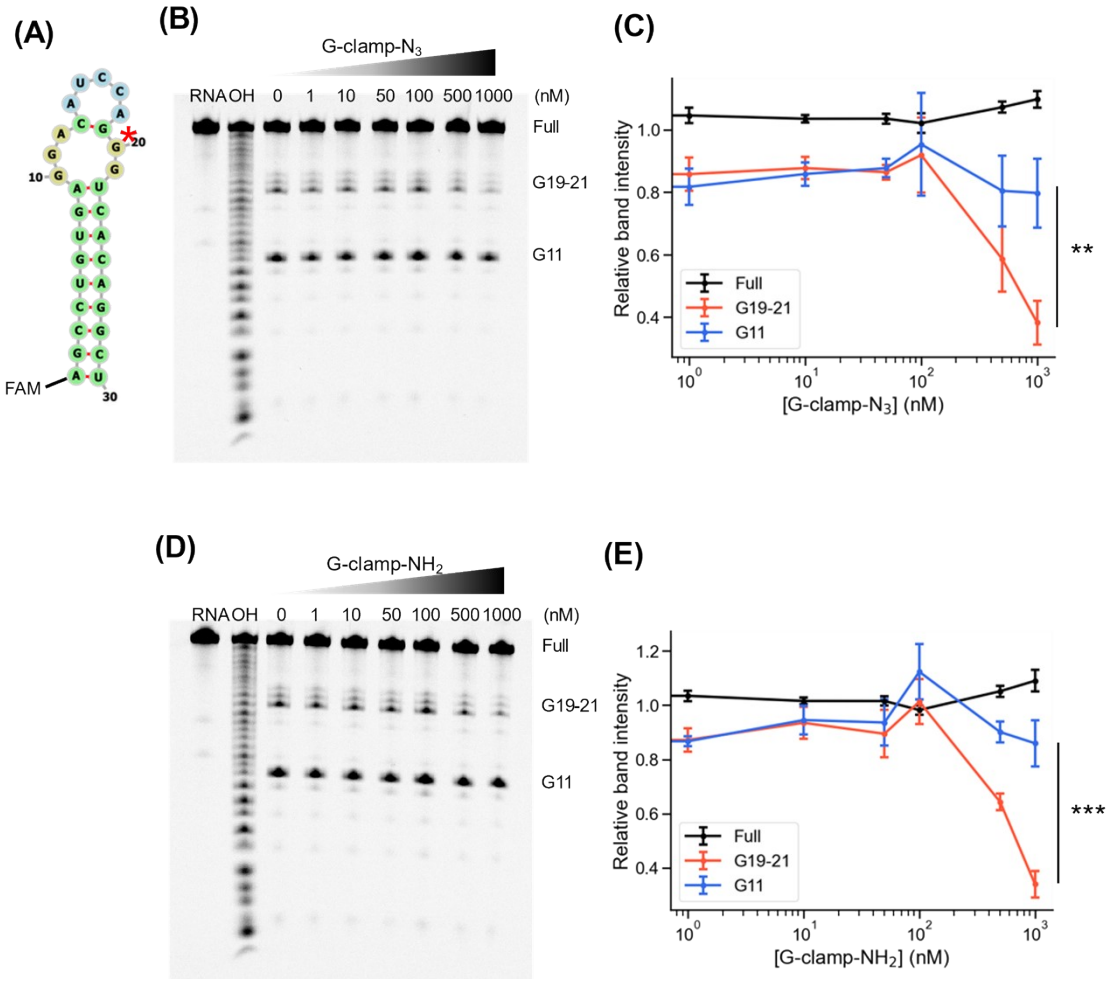


Figure S8. Footprinting assay using RNase T1. (**A**) Secondary structure of 5'FAM labeled pre-mir-125a predicted by RNAfold. Red star indicates the binding site of G-clamp derivatives. (**B**, **D**) Representative gel image of the inhibition of RNase T1 cleavage by (**B**) G-clamp-N_3 and (**D**) G-clamp-NH_2. (**C**, **E**) Relative band intensity of (**C**) G-clamp-N_3 and (**E**) G-clamp-NH_2. The data were mean ± SD obtained three independent experiments. *p*-values were calculated by two-tailed Student's T test. $^{**}p < 0.01$, $^{***}p < 0.001$.

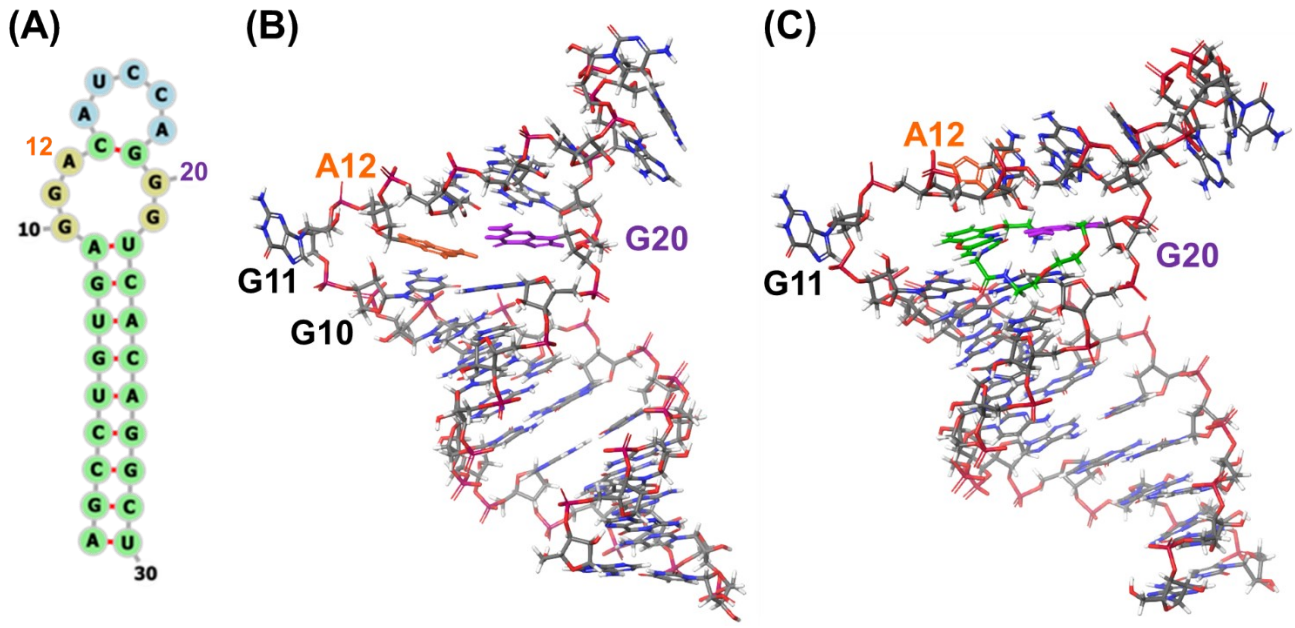


Figure S9. Molecular modeling of the complex structure between pre-mir-125a (WT) and G-clamp. (**A**) Secondary structure of pre-mir-125a (WT) predicted by RNAfold (-T 25). (B) Molecular modeling of pre-mir-125a (WT). The structure was modeled by RNAComposer and MacroModel. (C) Molecular modeling of the complex structure between pre-mir-125a (WT) and G-clamp. The green molecule indicates G-clamp. The orange and purple bases indicate A12 and G20, respectively. G-clamp binds to G20 and A12 is forced out from the inside.

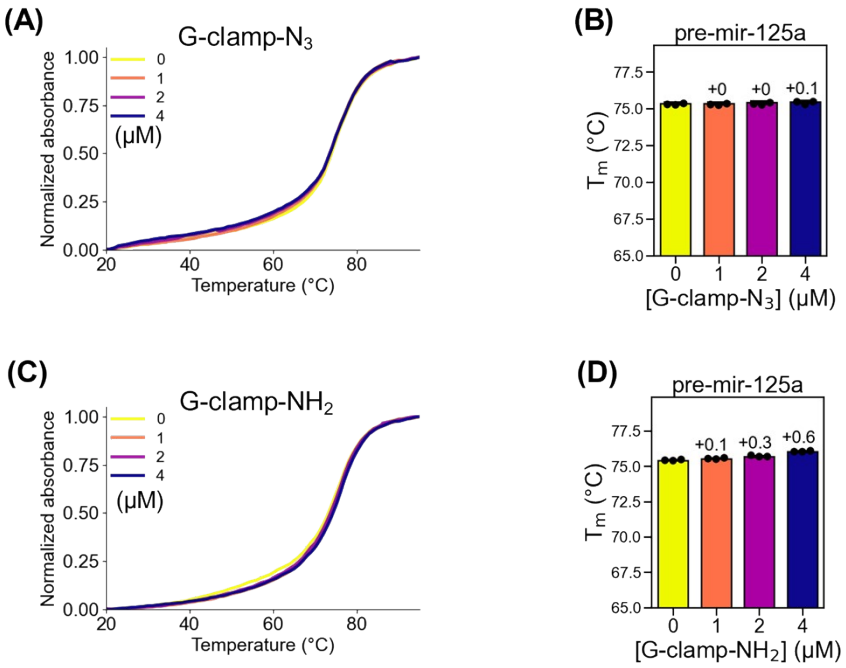
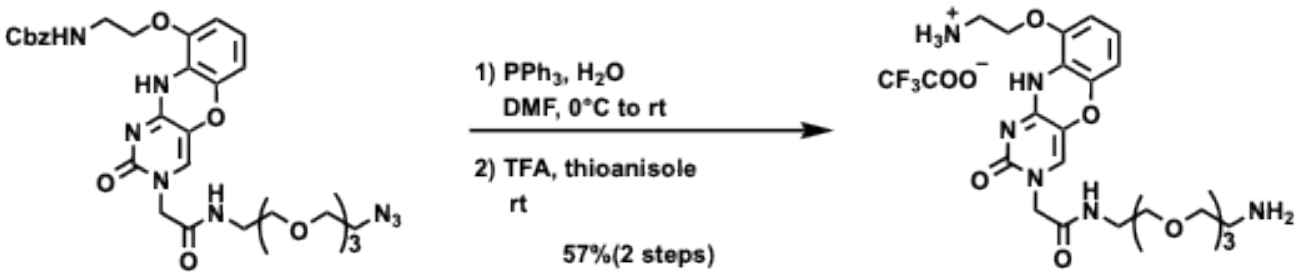


Figure S10. T_m analysis of pre-mir-125a with G-clamp derivatives. (**A**, **C**) Representative normalized T_m curves of pre-mir-125a (1 μM) with (**A**) G-clamp-N_3 or (**C**) G-clamp-NH_2 (0, 1, 2, and 4 μM) in 1x Binding buffer (1% DMSO, 20 mM phosphate pH 7.0, 20 mM NaCl, and 80 mM KCl). (**B**, **D**) Bar graphs of T_m values of pre-mir-125a with (**B**) G-clamp-N_3 or (**D**) G-clamp-NH_2. The data were mean ± SE obtained by three technical replicates. The number on the bar graphs represent delta T_m values (℃).

Synthesis of G-clamp-NH_2



1 **2**

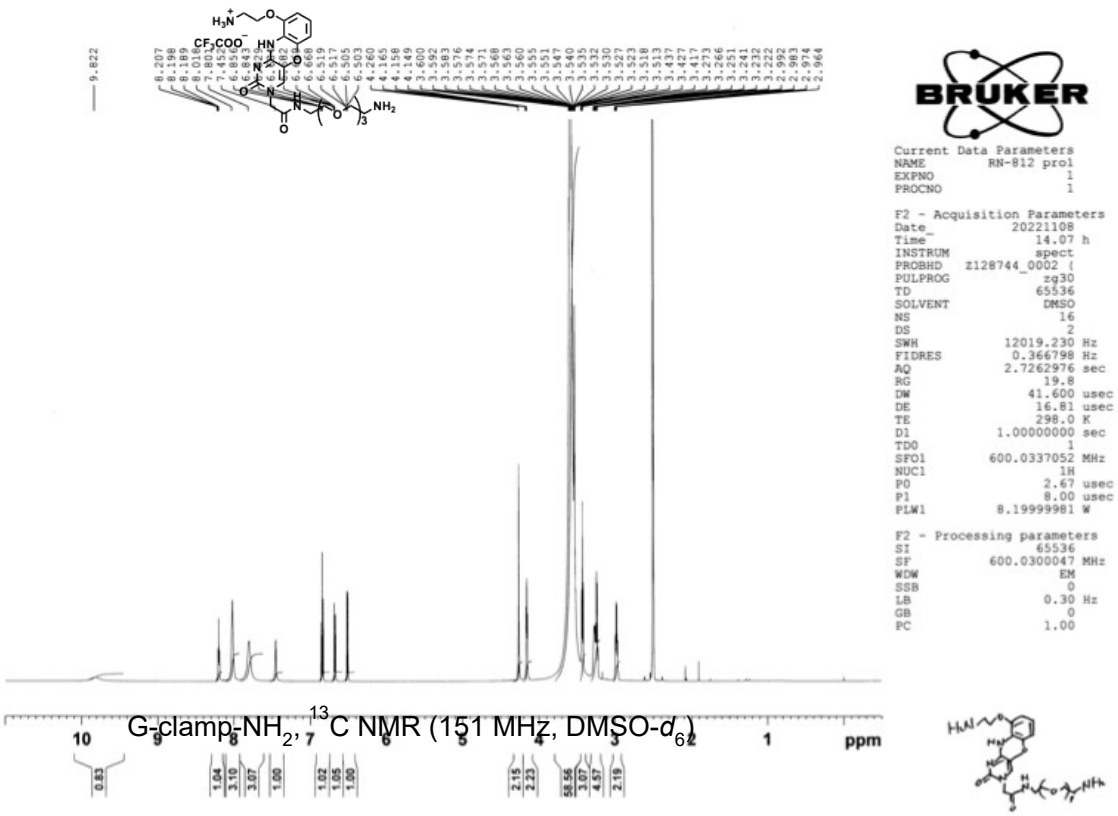
To a solution of the compound **1** (11.3 mg, 0.0173 mmol) in DMF (1 mL), PPh_3 (11 mg, 0.0419 mmol) was added at 0℃. After 5 min, the mixture was warmed to room temperature. After further 10 min, H_2O (0.2 mL) was added to the mixture and stirred for further 18 h. The mixture was washed with DCM (10 mL×2), H_2O (6 mL), and H_2O (10 mL×2) with AcOH (1 drop), and all aqueous layers were collected followed by evaporation under reduced pressure to afford the intermediate.

To a solution of the intermediate in TFA (0.2 mL), thioanisole (48 mg, 0.39 mmol) was added and stirred at room temperature for 18 h. Then, another TFA (0.2 mL) was added to the mixture and stirred for further 10 h. The mixture was evaporated under reduced pressure, filtrated with DMSO (1 mL), and purified by reverse phased HPLC to afford compound **2** (9.79 μmol, 57% (2 steps)) as a yellow solid.

HPLC conditions; A : 0.1% TFA in distilled water, B : 0.1% TFA in MeCN; B : 5%→60% (～20 min)→100%(～21 min)→100%(～26 min). Flow rate = 4 mL/min; Temp. =35.0 ℃; UV = 254 nm, C-18 column (Nacalai tesque : COSMOSIL 5C18-AR-Ⅱ, 10×250 mm).

^{1}H NMR (600 MHz, DMSO-d_6) δ 2.98 (2H, q, *J* = 5.6 Hz), 3.23 (2H, q, *J* = 5.7 Hz), 3.27 (2H, q, *J* = 4.2 Hz), 3.43 (2H, t, *J* = 6.0 Hz), 3.5-3.6 (10H, m), 4.16 (2H, t, *J* = 4.8), 4.26 (2H, s), 6.52 (1H, dd, *J* = 1.2, 9.6 Hz), 6.67 (1H, dd, *J* = 1.2, 9.0 Hz), 6.84 (1H, t, *J* = 8.1 Hz), 7.45 (1H, s), 7.80 (3H, br s), 8.02 (3H, br s), 8.20 (1H, t, *J* = 5.4 Hz), 9.82 (1H, br s). ^{13}C NMR (151 MHz, DMSO-d_6) δ38.7, 50.7, 64.9, 66.7, 69.0, 69.58, 69.62, 69.67, 69.70, 107.0, 108.5, 114.3, 115.7, 116.3, 118.3, 120.3, 123.2, 125.6, 129.8, 142.3, 145.4, 154.2, 157.5, 157.7, 157.9, 158.1, 167.2. ESI-HRMS (m/z): $[M+H]^+$ calcd for $C_{22}H_{33}N_6O_7^+$, 493.2405, found 493.2410.

G-clamp-NH_2, ^{1}H NMR (600 MHz, DMSO-d_6)

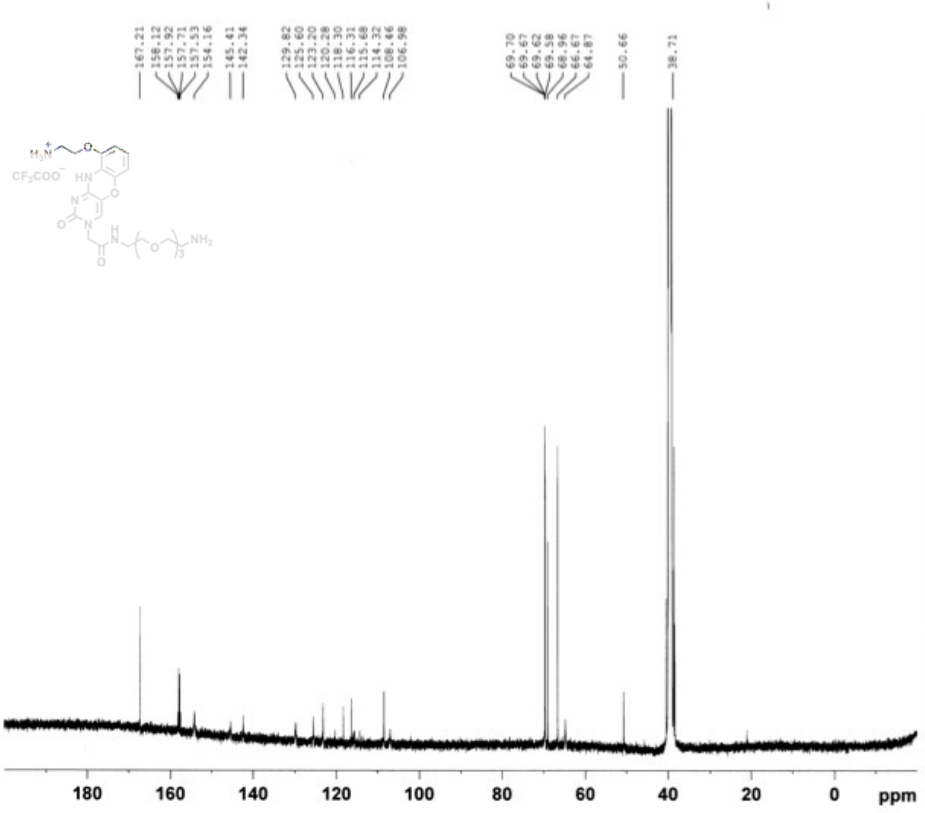



```
Current Data Parameters
NAME          RN-812 prol
EXPNO                   1
PROCNO                  1

F2 - Acquisition Parameters
Date_            20221108
Time                14.07 h
INSTRUM              spect
PROBHD    Z128744_0002 (
PULPROG              zg30
TD                  65536
SOLVENT              DMSO
NS                     16
DS                      2
SWH           12019.230 Hz
FIDRES         0.366798 Hz
AQ             2.7262976 sec
RG                  19.8
DW                41.600 usec
DE                16.81 usec
TE                298.0 K
D1            1.00000000 sec
TD0                     1
SFO1         600.0337052 MHz
NUC1                  1H
P0                 2.67 usec
P1                 8.00 usec
PLW1          8.19999981 W

F2 - Processing parameters
SI                  65536
SF           600.0300047 MHz
WDW                   EM
SSB                    0
LB                  0.30 Hz
GB                     0
PC                  1.00
```

G-clamp-NH_2, ^{13}C NMR (151 MHz, DMSO-d_6)



BRUKER

```
Current Data Parameters
NAME      RN-812 prol 13C-2
EXPNO                    1
PROCNO                   1

F2 - Acquisition Parameters
Date_              20240418
Time                   9.42 h
INSTRUM               spect
PROBHD   Z128744_0002 (
PULPROG             zgpg30
TD                   65536
SOLVENT                DMSO
NS                   15682
DS                       4
SWH              36231.883 Hz
FIDRES            1.105709 Hz
AQ                0.9043968 sec
RG                   198.89
DW                   13.800 usec
DE                   18.00 usec
TE                   298.0 K
D1                2.00000000 sec
D11               0.03000000 sec
TD0                      1
SFO1           150.8927513 MHz
NUC1                    13C
P0                    4.00 usec
P1                   12.00 usec
PLW1             90.00000000 W
SFO2            600.0324001 MHz
NUC2                    1H
CPDPRG[2            waltz65
PCPD2                70.00 usec
PLW2              2.33240008 W
PLW12             0.10710000 W
PLW13             0.05387100 W

F2 - Processing parameters
SI                     32768
SF             150.8777350 MHz
WDW                       EM
SSB                        0
LB                    1.00 Hz
GB                         0
PC                    1.40
```

Reference

[1] P. Kerpedjiev, S. Hammer, I. L. Hofacker, *Bioinformatics*, 2015, 31, 3377–3379.

[2] R. Lorenz, S. H. Bernhart, C. Höner Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker, *Algorithms Mol. Biol*. 2011, 6, 26.

[3] F. H. Stootman, D. M. Fisher, A. Rodger, J. R. Aldrich-Wright, *Analyst*, 2006, 131, 1145–1151.

[4] W. Kabsch. *Acta Crystallogr. D Biol. Crystallogr*. 2010, 66, 125-132.

[5] P. D Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R, Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr*. 2010, 66, 213-221.

[6] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* 2007, 40, 658-674.

[7] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr*. 2002, 60, 2126-2162.

[8] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D Biol*. Crystallogr. 2010, 66, 486-501.

[9] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr*. 2012, 68, 352-367.

[10] W. L. DeLano, The PyMOL Molecular Graphics System. DeLano Scientific LLC, Palo Alto, CA, USA. 2008.

[11] R. Nagasawa, K. Onizuka, K. R. Komatsu, E. Miyashita, H. Murase, K. Ojima, S. Ishikawa, M. Ozawa, H. Saito and F. Nagatsugi, *Commun Chem*, 2024, 7, 98.