Supporting Information

A Preorganized β-Amino Acid Bearing a Guanidinium Side Chain and its Use in Cell-Penetrating Peptides

Yosuke Demizu,*[a],[b] Makoto Oba,[c] Kouyo Okitsu,[b],[d] Hiroko Yamashita,[b],[d]

Takashi Misawa,[b] Masakazu Tanaka,[c] Masaaki Kurihara,[b],[d]

and Samuel H. Gellman[a]

[a]*Departments of Chemistry, University of Wisconsin, Madison, 53706, USA;* [b]*Division of Organic Chemistry, National Institute of Health Sciences, Tokyo 158-8501, Japan;* [c]*Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8521, Japan;* [d]*Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan*

masaaki@nihs.go.jp

Supporting Information

General. ^{1}H and ^{13}C NMR spectra were recorded at 400 MHz and 100 MHz in $CDCl_3$ with tetramethylsilane used as an internal standard. CD spectra were recorded with a *Jasco J-720W* spectropolarimeter using a 1.0 mm path length cell. The data were expressed in terms of $[\theta]_R$, the total molar ellipticity (deg cm^2 $dmol^{-1}$). Methanol was used as a solvent.

Cellular uptake. HeLa cells were seeded onto 96-well culture plates (10,000 cells/well) and incubated overnight in 100 μL of DMEM containing 10% FBS. The medium was replaced by fresh medium with 10% FBS and Tat-peptide solution was applied to each well to become a concentration of 1 μM. After 2h incubation, the medium was removed, and the cells were washed 3 times with ice-cold PBS supplemented with heparin (20 units/mL) and treated with Cell lysis buffer M. The fluorescence intensity of lysate was measured by a spectrofluorometer (ND-3300, NanoDrop, Wilmington, DE, USA). The amount of protein in each well was concomitantly determined using a Micro BCA protein assay reagent kit. The results are presented as the mean and standard error obtained from 5 samples.

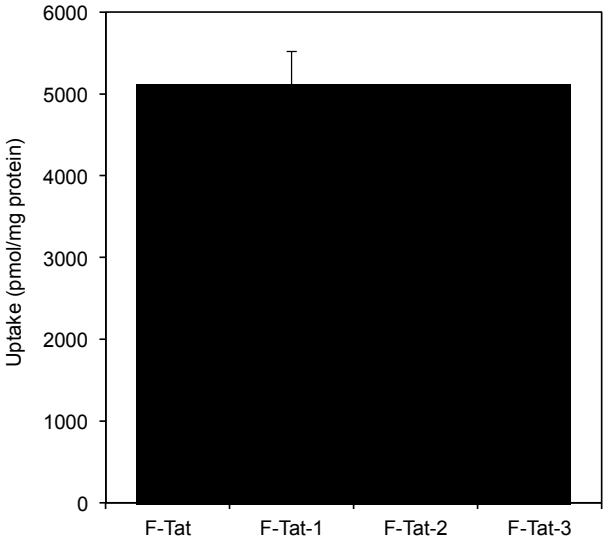


Fig. S1. Cellular uptake of **F-Tat**, **F-Tat-1-3** at concentrations of 2 μM (incubation time: 2 hr). Error bars represent standard deviation, n = 3.

Inhibition of endocytosis. HeLa cells were seeded onto 96-well culture plates (10,000 cells/well) and incubated overnight in 100 μL of DMEM containing 10% FBS. After replacement with fresh medium containing 10% FBS in the absence or presence of amiloride (5 mM), chlorpromazine (10 μg/mL), and filipin (5 μg/mL), cells were pre-incubated at 37˚C or 4˚C for 30 min. Tat-peptide solution was applied to each well to become a concentration of 1 μM. After 2h incubation, the medium was removed, and the cells were washed 3 times with ice-cold PBS supplemented with heparin (20 units/mL) and treated with Cell lysis buffer M. The fluorescence intensity of lysate was measured by a spectrofluorometer (ND-3300, NanoDrop, Wilmington, DE, USA). The amount of protein in each well was concomitantly determined using a Micro BCA protein assay reagent kit. The results are presented as the mean and standard error obtained from 4 samples.

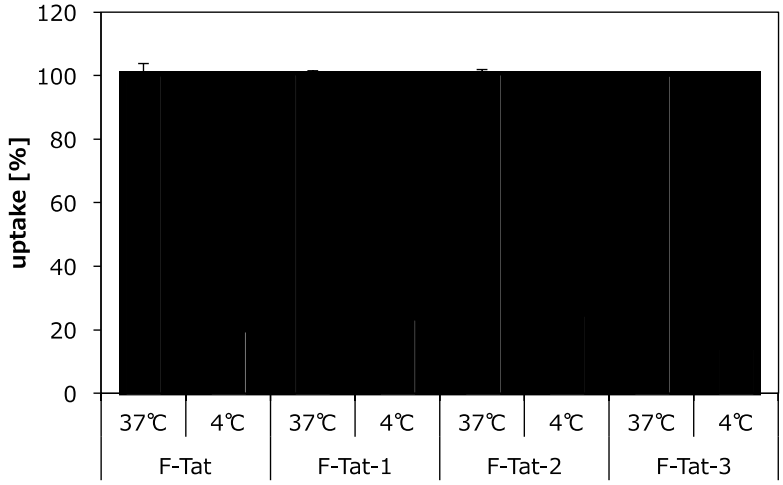


Fig. S2. Effects of 4˚C incubation on internalization of **F-Tat**, **F-Tat-1-3** at concentrations of 1 μM (incubation time: 2 hr). Error bars represent standard deviation, n = 3.

CLSM observation. HeLa cells were seeded onto 8-well chambered coverglasses (Iwaki, Tokyo, Japan) (20,000 cells/well) and incubated overnight in 200 μL of DMEM containing 10% FBS. The medium was replaced by fresh medium with 10% FBS and Tat-peptide solution was applied to

each well to become a concentration of 1 μM. After 2h incubation, the medium was removed, and the cells were washed 3 times with ice-cold PBS supplemented with heparin (20 units/mL). The intracellular distribution of the complexes was observed by CLSM after staining late endosomes/lysosomes with LysoTracker Red and nuclei with Hoechst 33342. The CLSM observation was performed using an LSM 710 (Carl Zeiss, Oberlochen, Germany) equipped with a 63X objective (Plan-Apochromat, Carl Zeiss) or a 100X objective (Plan-Apochromat, Carl Zeiss) at excitation wavelengths of 405 nm (UV laser) for Hoechst 33342, 488 nm (Ar laser) for Tat, and 543 nm (He-Ne laser) for LysoTracker Red, respectively. The rate of colocalization of peptides with LysoTracker Red was quantified. Colocalization was quantified as follows: colocalization ratio (%) = Peptide pixels colocalization/Peptide pixels total X 100, where peptide pixels colocalization represents the number of peptide pixels colocalizing with LysoTracker Red in the cell, and peptide pixels total represents the number of all the peptide pixels in the cell. The results are presented as the mean and standard error obtained from 15 cells.

CD spectra of Tat peptides.

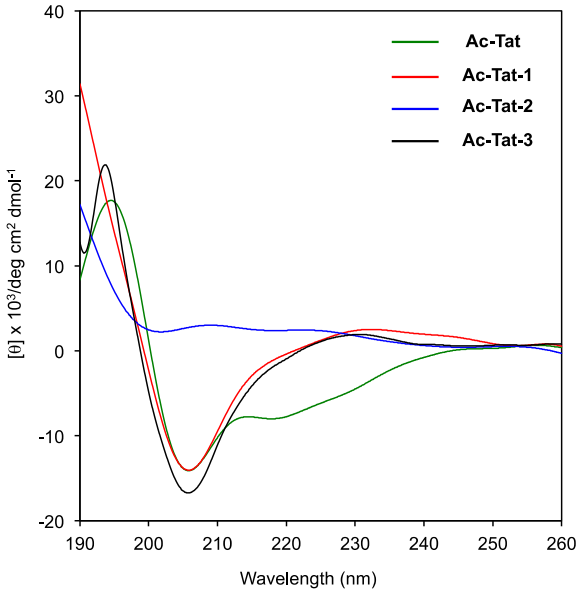


Fig. S3 CD spectra in the 190—260 nm region of **Ac-Tat** (green), **Ac-Tat-1** (red), **Ac-Tat-2** (blue), and **Ac-Tat-3** (black). Peptide concentration: 50 μM in 10mM SDS/MeOH solution.

Synthesis of Fmoc-(3*S*,4*R*)-APC^{Gu}-OH (3). A solution of Fmoc-(3*S*,4*R*)-APC^{Boc}-OH (**1**) (453 mg, 1.0 mmol) in 4M HCl/1,4-dioxane (3 mL) was stirred at room temperature for 2 h. Removal of the solvent afforded a crude Fmoc-(3*S*,4*R*)-APC-OH, which was used without further purification. The above amino acid (1.0 mmol) was suspended in CH_2Cl_2, and N-methyl-N-trimethylsilyl-trifluoroacetamide (222 μL, 1.2 mmol) was added into the solution. The mixture was refluxed for 20 min to obtain a clear solution. After cooling to room temperature, N,N'-di-Boc-N''-triflylguanidine (1.17 g, 3.0 mmol) and Et_3N (418 μL, 3.0 mmol) were added into the solution, and the mixture was stirred at room temperature for 12h. The mixture was washed with 2M sodium bisulfate and water, and dried over $MgSO_4$. After removal of the solvent, the residue was purified by column chromatography on silica gel (0.1% AcOH in AcOEt) to give Fmoc-(3*S*,4*R*)-APC^{Gu}-OH (**3**) (565 mg, 95%). Colorless crystals; Mp 109—112°C; ^{1}H NMR (400 MHz, DMSO-d6) δ 9.52 (s, 1H), 7.87 (d, *J* = 6.0 Hz, 2H), 7.67 (d, *J* = 6.0 Hz, 2H), 7.61 (br s, 1H), 7.39 (t, *J* = 6.0 Hz, 2H), 7.31 (t, *J* = 6.0 Hz, 2H), 4.17—4.30 (m, 4H), 3.58—3.64 (m, 3H), 3.21 (br s, 1H), 2.76 (br, 1H), 1.39 (s, 9H), 1.38 (s, 9H); ^{13}C NMR(100 MHz, $CDCl_3$) δ 174.1, 156.1, 153.3, 143.7, 143.6, 141.2, 127.6, 127.1, 125.1, 119.9, 81.4, 67.0, 53.2, 52.9, 49.4, 48.2, 46.9, 28.4, 28.0, 27.9; [HR-ESI-TOF(+)]: *m/z* calcd for $C_{31}H_{39}N_4O_8$ $[M+H]^+$ 595.2768: found 595.2771.

Synthesis of Tat-peptides. The peptides were synthesized using Fmoc-solid phase methods on NovaPEG Rink amide resin. Microwave irradiation was used for the coupling-deprotection cycle. The following describes a representative coupling and deprotection cycle at a 25 μmol scale. First, 65 mg NovaPEG Rink amide resin (loading: 0.5 mmol/g) were soaked for 1 hr in CH_2Cl_2. After the resin had been washed with N-methyl-2-pyrrolidone (NMP), Fmoc-amino acid (4 Eq), and HBTU (4 Eq) dissolved in 1.5 mL NMP were added to the resin. Then, *N*,*N*-diisopropylethylamine (4 Eq) and hydroxybenzotriazole (1.0 mL, 0.1 M solution in NMP) were added for the coupling reaction. Deprotection was carried out using 20% piperidine in dimethylformamide (2 mL). After the peptide synthesis, the resin was suspended in cleavage cocktail (1.9 mL trifluoroacetic acid [TFA], 50 μL water, 50 μL triisopropylsilane; final concentration: 95% TFA, 2.5% water, 2.5% triisopropylsilane) for 3 hr at rt. The TFA solution was evaporated to a small volume under a stream of N_2 and dripped into cold ether to precipitate the peptides. The dried crude peptides were dissolved in 2 mL of 50% acetonitrile in water and then purified by reversed-phase HPLC using a Discovery® BIO Wide Pore C18 column (25 cm x 21.2 mm). After being purified, the peptide solutions were lyophilized. Peptide purity was assessed using analytical HPLC, and were characterized by MALDI-TOF mass spectrometry.

Purities of **F-Tat** and **F-Tat-1-3** peptides were analyzed by reverse-phase HPLC using VYDAC® protein & peptide C18 column (218TP54) (25 cm x 4.6 mm)

solvent A; 0.1% TFA/water

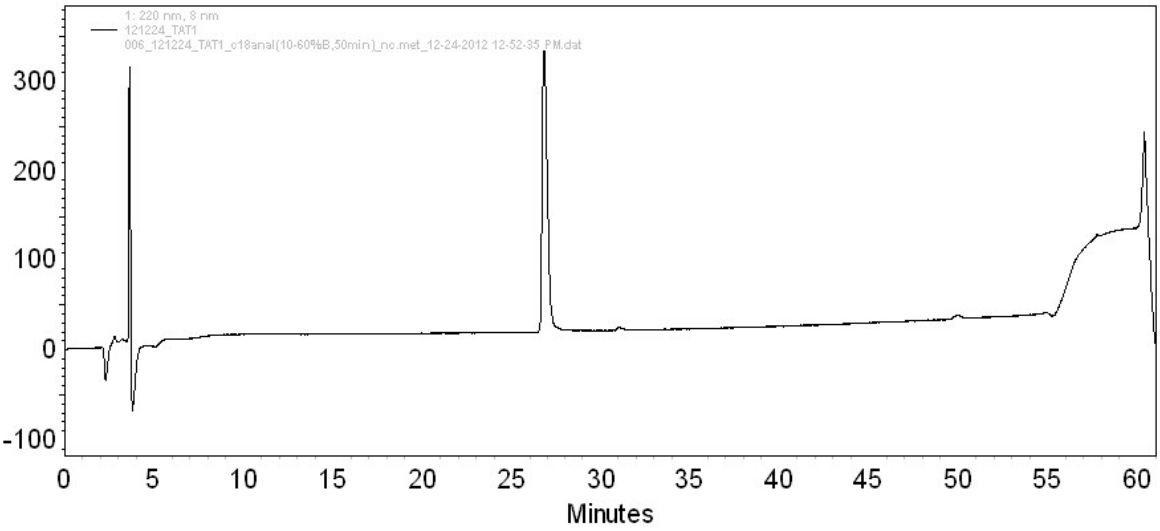
solvent B; 0.1% TFA/MeCN

flow rate, 1.0 mL•mL^{-1}

gradient, 10-60% gradient of solvent B over 50 min

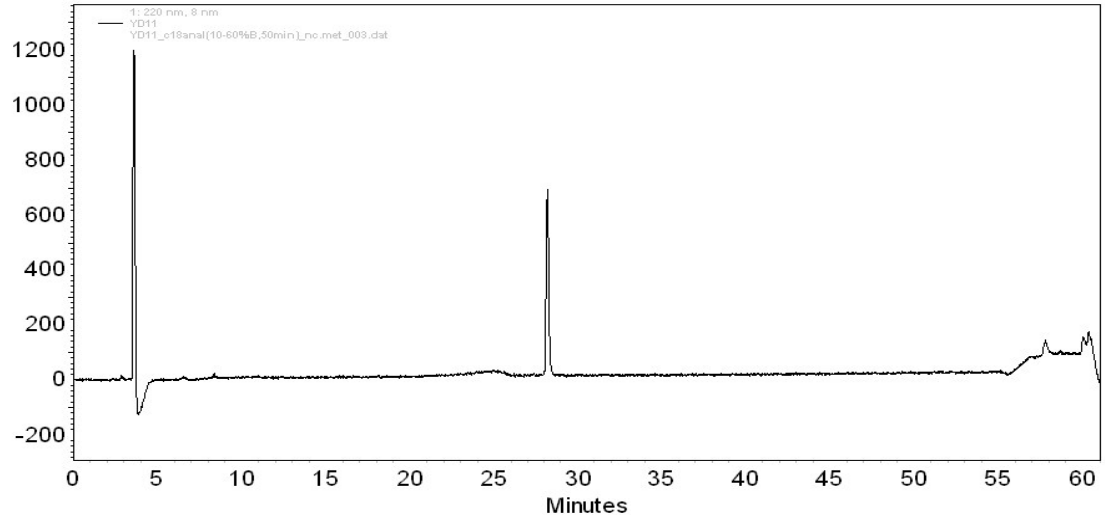
F-Tat

MALDI-TOF(+)-MS *m/z* 1989.0 $(M+H)^+$



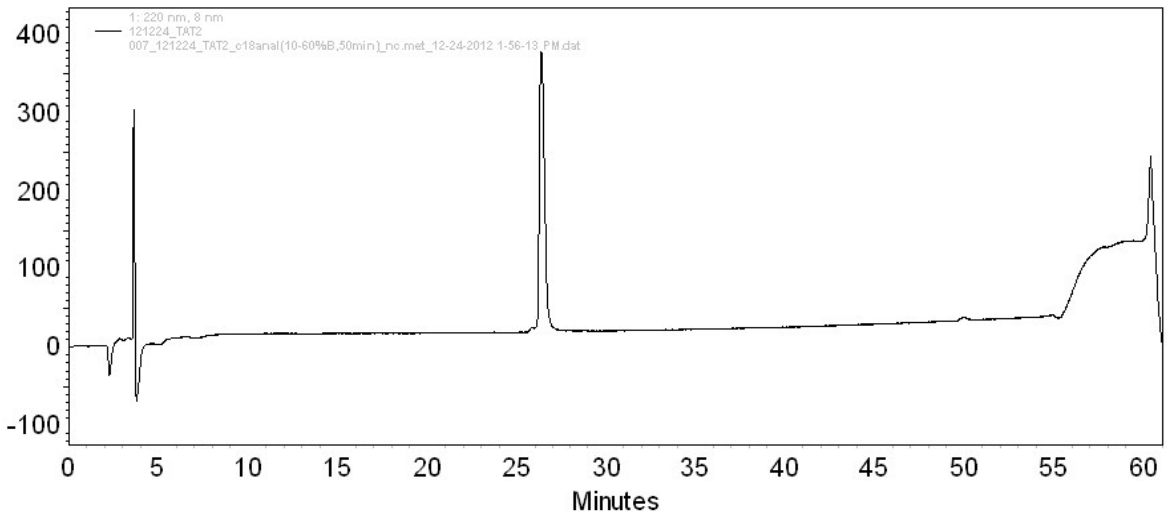
F-Tat-1

MALDI-TOF(+)-MS *m/z* 1981.3 $(M)^+$



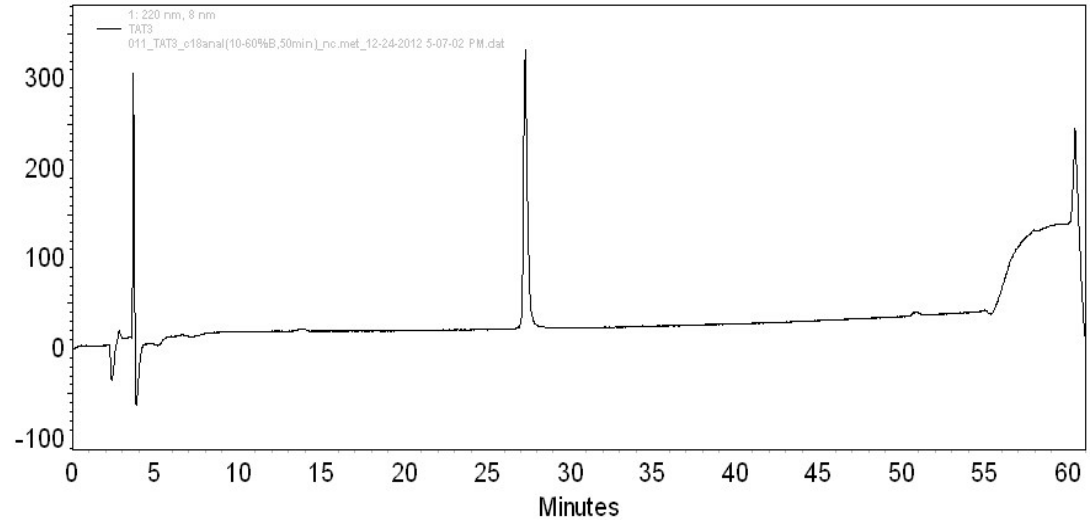
F-Tat-2

MALDI-TOF(+)-MS *m/z* 2030.8 $(M+H)^+$

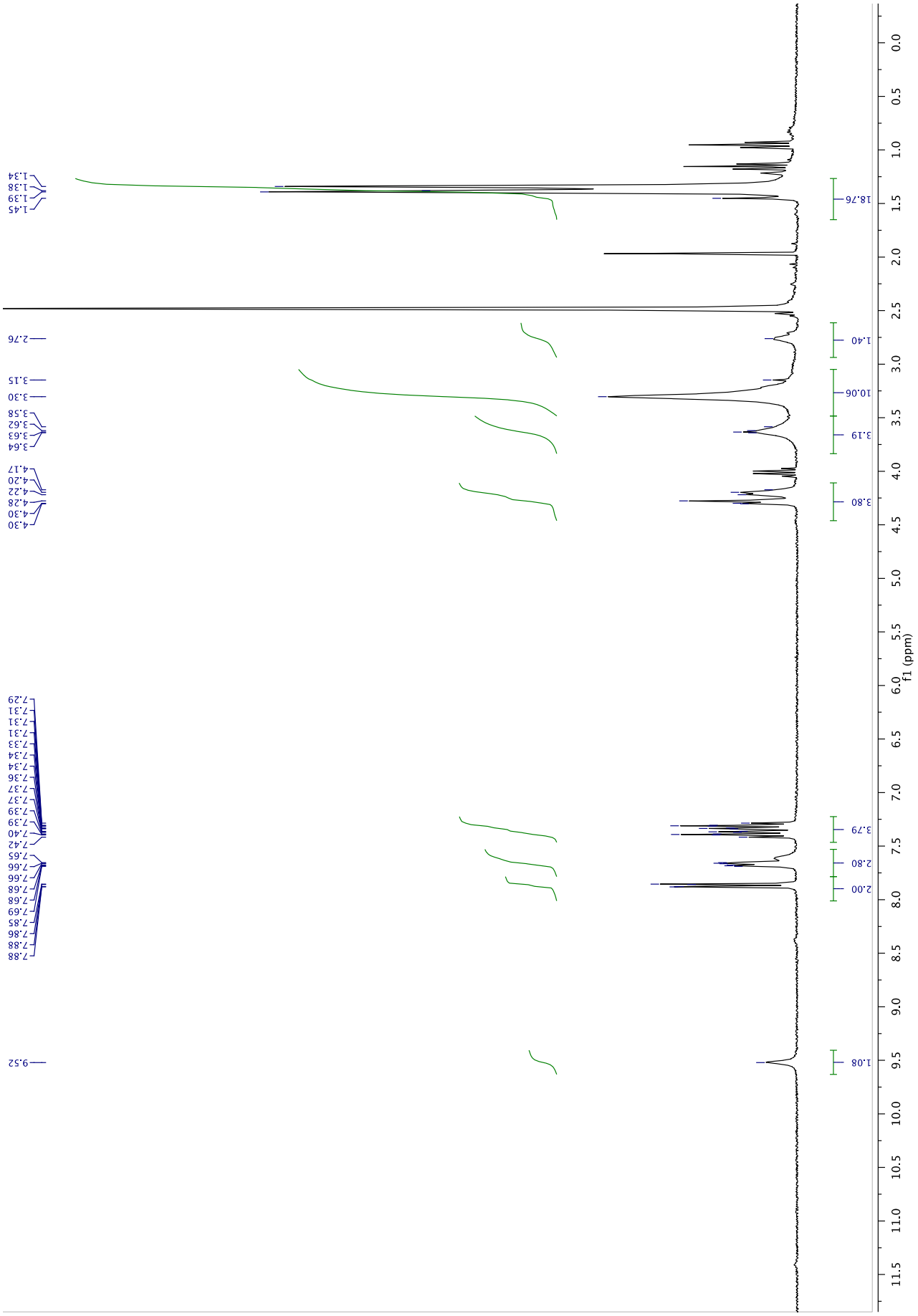


F-Tat-3

MALDI-TOF(+)-MS *m/z* 1856.0 $(M+H)^+$



^{1}H spectrum of Fmoc-(3*S*,4*R*)-APC^{Gu}-OH (**3**)



^{13}C spectrum of Fmoc-(3*S*,4*R*)-APC^{Gu}-OH (**3**)