Supporting Information

Evaluation of cationic assemblies constructed with amino-acid based lipids for plasmid DNA delivery

Yosuke Obata, Daisuke Suzuki, and Shinji Takeoka

Department of Life Science and Medical Bioscience, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo 169-8555, Japan.

y-obata@asagi.waseda.jp

List of contents: **Page**

Synthesis of the hydrophobic moiety of cationic lipids. a (1,5-ditetradecyl-L-glutamate), **b** (1,5-dihexadecyl-L-glutamate), and **c** (1,5-dioctadecyl-L-glutamate) were synthesized as shown in Scheme. L-Glutamic acid (1 g, 6.8 mmol) and *p*-Tos (1.65 g, 8.16 mmol) were dissolved in benzene (200 mL) and refluxed for 1 hr at 90°C. Tetradecylalcohol, hexadecylalcohol or octadecylalcohol (15 mmol) was added to the solution, followed by stirring for 12 hr under reflux. The reaction mixture was evaporated and then dissolved in chloroform (100 ml). The chloroform solution was treated with a sodium carbonate solution (100 ml x 2) and washed with distilled water (100 ml x 1). After the chloroform solution was evaporated, **a**, **b** and **c** were recrystalized from methanol (100 mL) at 4°C to obtain a white powder with a yield of 76%, 82% and 83 %, respectively.

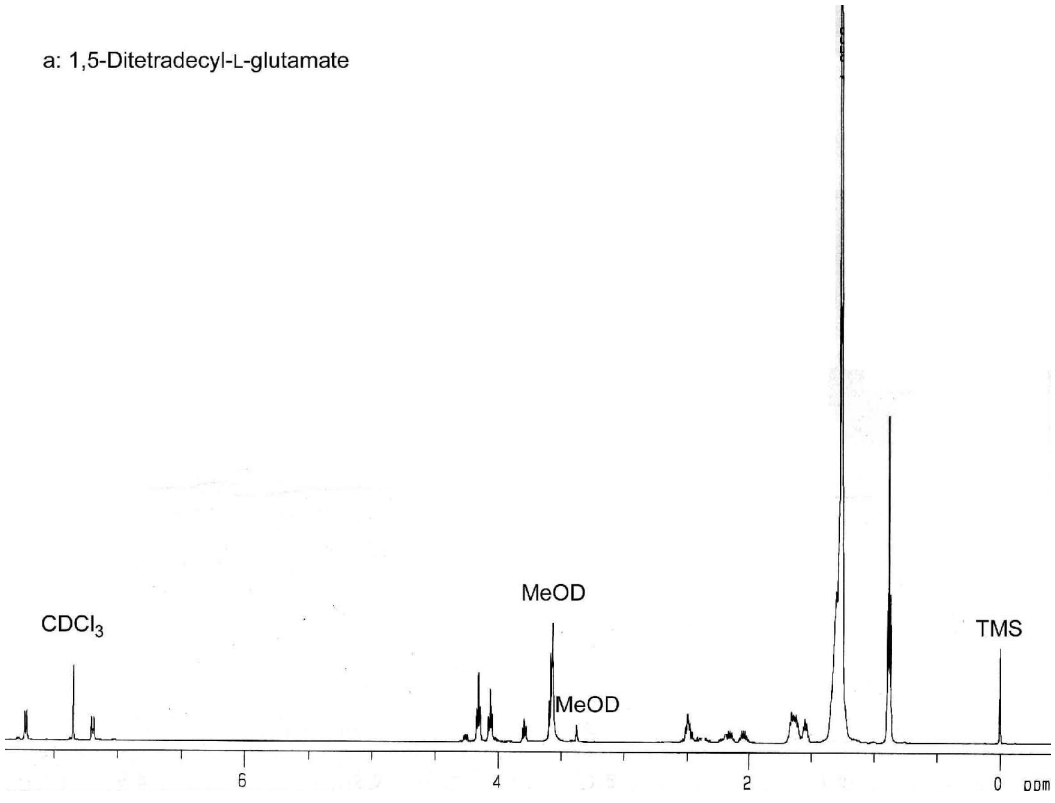

Figure 1S. ^{1}H-NMR spectrum of 1,5-ditetradecyl-L-glutamate (**a**).

1,5-Ditetradecyl-L-glutamate (a): R_f 0.75 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.28 (m, 44H, $CH_{2(myristoyl)}$), 1.58-1.66 (m,4H, $COOCH_2CH_2$); 2.01-2.10 (m, 2H, NH_2CHCH_2), 2.45 (t, 2H, CH_2CO); 3.80 (t, 1H, NH_2CH), 4.06-4.14 (t, 4H, $COOCH_2$), 7.18, 7.73 (d, 2H, NH_2). MS(ESI): $(M+H)^+$ calcd. for $C_{33}H_{65}NO_4$, 539.87; found, 540.5.

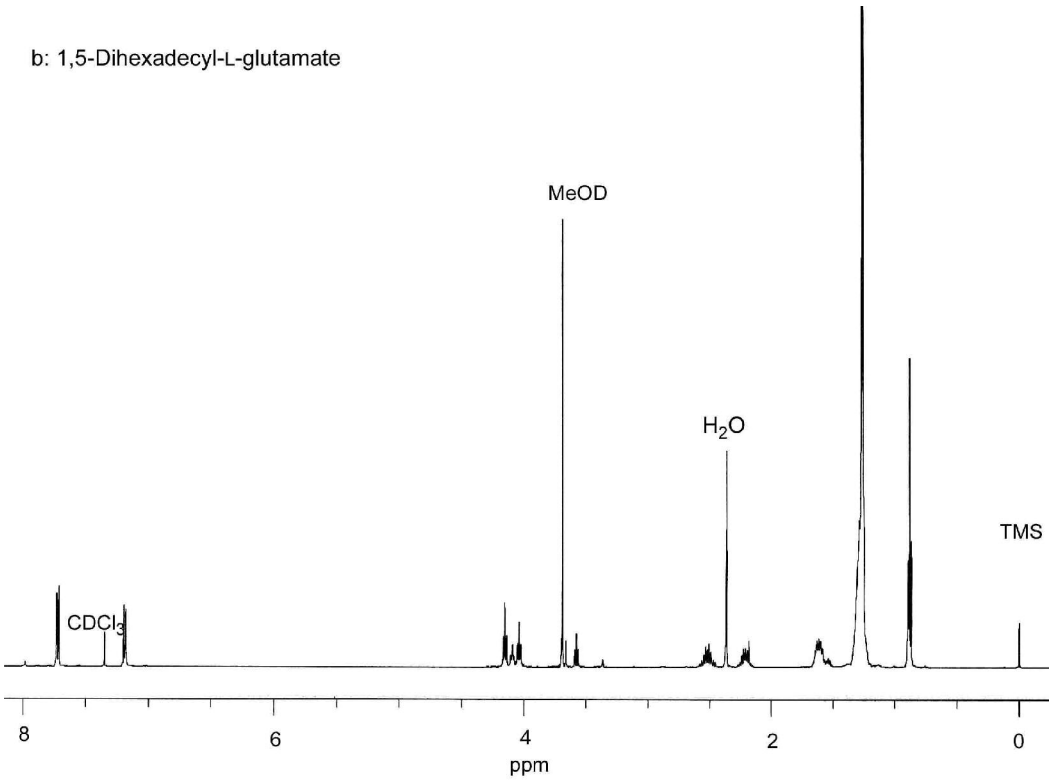

Figure 2S. ^{1}H-NMR spectrum of 1,5-dihexadecyl-L-glutamate (**b**).

1,5-Dihexadecyl-L-glutamate (b): R_f 0.73 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 52H, $CH_{2(palmitoyl)}$), 1.55-1.64 (m,4H, $COOCH_2CH_2$); 2.21 (m, 2H, NH_2CHCH_2), 2.51 (t, 2H, CH_2CO), 4.04 (t, 1H, NH_2CH), 4.10-4.14 (t, 4H, $COOCH_2$), 7.18, 7.72 (d, 2H, NH_2). MS(ESI): $(M+H)^+$ calcd. for $C_{37}H_{73}NO_4$, 595.58; found, 595.5.

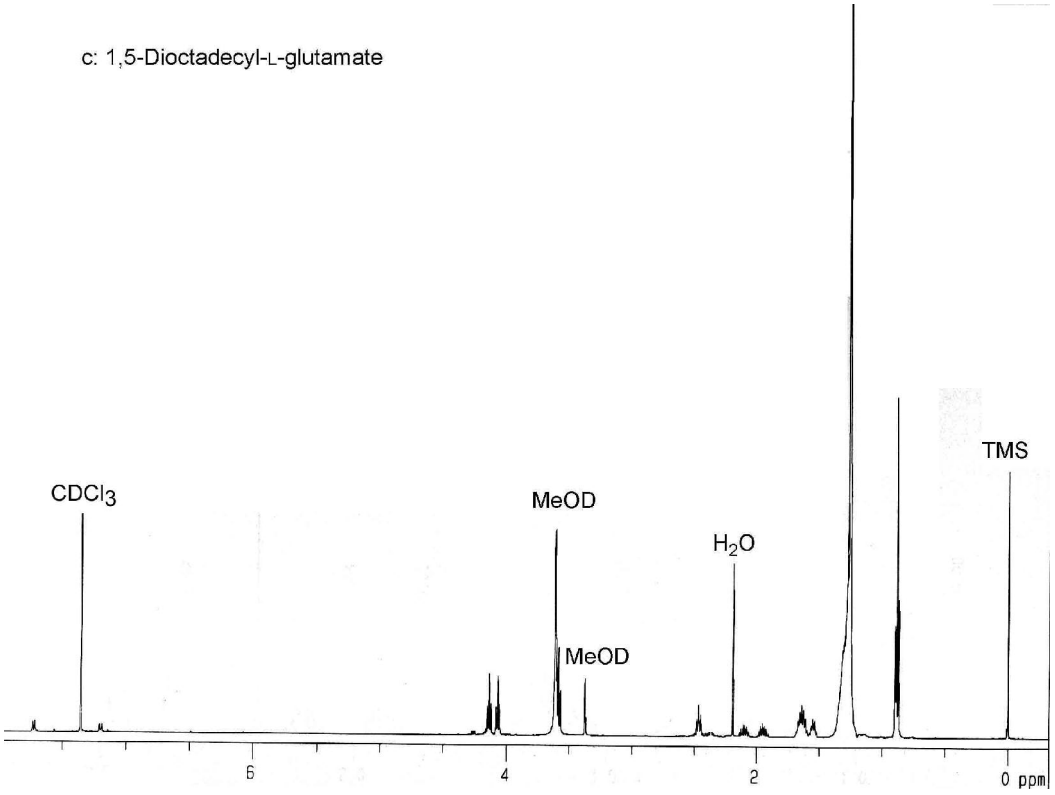

Figure 3S. ^{1}H-NMR spectrum of 1,5-dihexadecyl-L-glutamate (**c**).

1,5-Dihexadecyl-L-glutamate (c): R_f 0.75 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 60H, $CH_{2(stearyl)}$), 1.52-1.67 (m,4H, $COOCH_2CH_2$), 1.90-2.12 (m, 2H, NH_2CHCH_2), 2.47 (t, 2H, CH_2CO), 3.59 (t, 1H, NH_2CH), 4.06-4.15 (t, 4H, $COOCH_2$), 7.18, 7.73 (d, 2H, NH_2). MS(ESI): $(M+H)^+$ calcd for $C_{33}H_{65}NO_4$, 652.09; found, 652.5.

Syntheses of amino-acid based cationic lipids. Amino acid-based cationic lipids having a cationic head-group were synthesized as shown in Scheme. **a**, **b** or **c** (1.68 mmol) and triethylamine (2 mmol) were dissolved in dichloromethane and stirred for 1 hr at room temperature. Either Boc-Lys(Boc)-OSu, Boc-His(1-Boc)-OSu, or Boc-Arg(Boc)$_2$-OH was added to the solution and stirred for 6 hr at room temperature. The reaction mixture was evaporated and then dissolved in chloroform (100 ml). The chloroform solution was treated with a sodium carbonate solution (100 ml x 2) and washed with distilled water (100 ml x 1). After the chloroform solution was evaporated, the amino group-protected intermediates were recrystalized from methanol (50 mL) at 4^{o}C. After deprotection with trifluoroacetic acid (20 mL) for 2 hr at 4^{o}C, **1a** (1,5-ditetradecyl *N*-lysyl-L-glutamate), **1b** (1,5-dihexadecyl *N*-lysyl-L-glutamate) or **1c** (1,5-dioctadecyl *N*-lysyl-L-glutamate) were obtained as a white powder (yields: 61%, 90%, or 86%, respectively) after freeze-drying with benzene. Using a similar method to that described above, **2a** (1,5-ditetradecyl *N*-histidyl-L-glutamate), **2b** (1,5-dihexadecyl *N*-histidyl-L-glutamate) and **2c** (1,5-dioctadecyl *N*-histidyl-L-glutamate) were obtained with a yield of 68%, 81% and 80%, respectively. For the synthesis of arginine-type lipids, BOP reagent was used for the amide linkage. Compounds **3a** (1,5-ditetradecyl *N*-arginyl-L-glutamate), **3b** (1,5-dihexadecyl *N*-arginyl-L-glutamate) and **3c** (1,5-dioctadecyl *N*-arginyl-L-glutamate) were obtained with a yield of 40%, 55% and 70%, respectively.

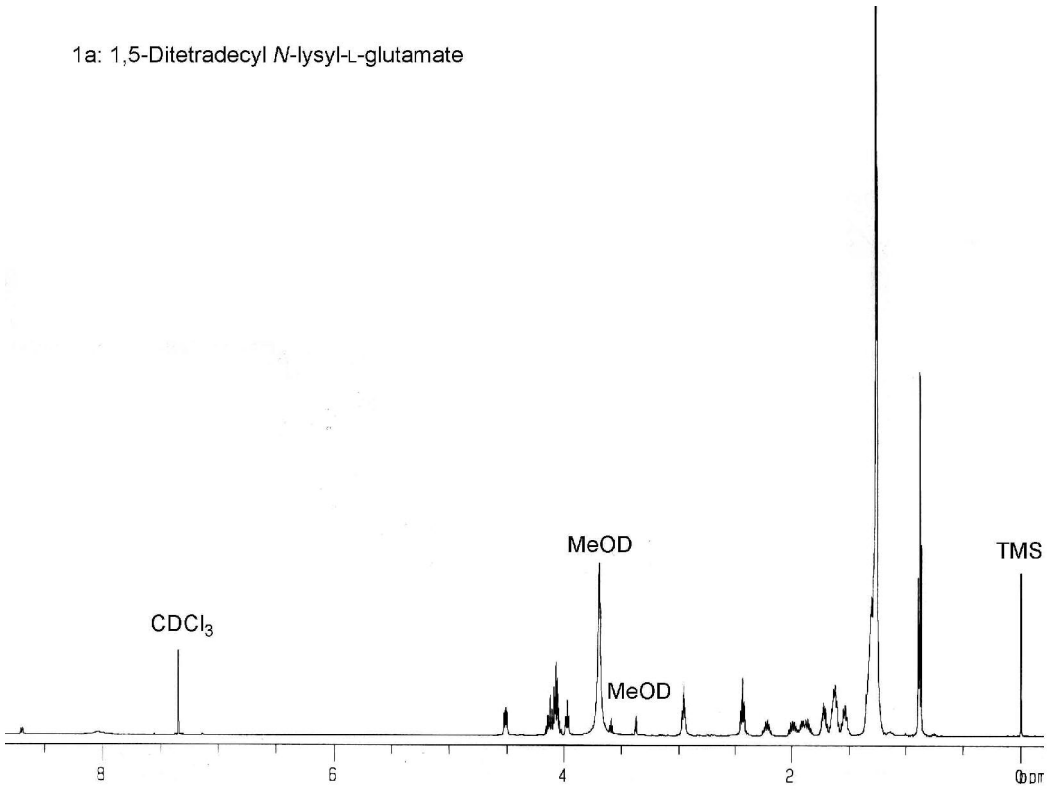

Figure 4S. ^{1}H-NMR spectrum of 1,5-ditetradecyl *N*-lysyl-L-glutamate (**1a**).

1,5-Ditetradecyl *N*-lysyl-L-glutamate (1a): R_f 0.14 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 44H, $CH_{2(myristoyl)}$), 1.53 (m, 2H, $NH_2CHCH_2CH_2$), 1.62 (m, 4H, $COOCH_2$), 1.72 (m, 2H, $NH_2CH_2CH_2$), 1.80-1.95 (m, 2H, NH_2CHCH_2), 1.96-2.25 (m, 2H, $NHCHCH_2$), 2.43 (m, 2H, CH_2COO), 2.96 (m, 2H, NH_2CH_2), 3.97 (m, 1H, NH_2CHCO), 4.00-4.20 (m, 4H, $COOCH_2CH_2$), 4.51 (q, 1H, NHCH). MS(ESI): $(M+H)^+$ calcd. for $C_{39}H_{77}N_3O_5$, 668.05; found, 669.7.

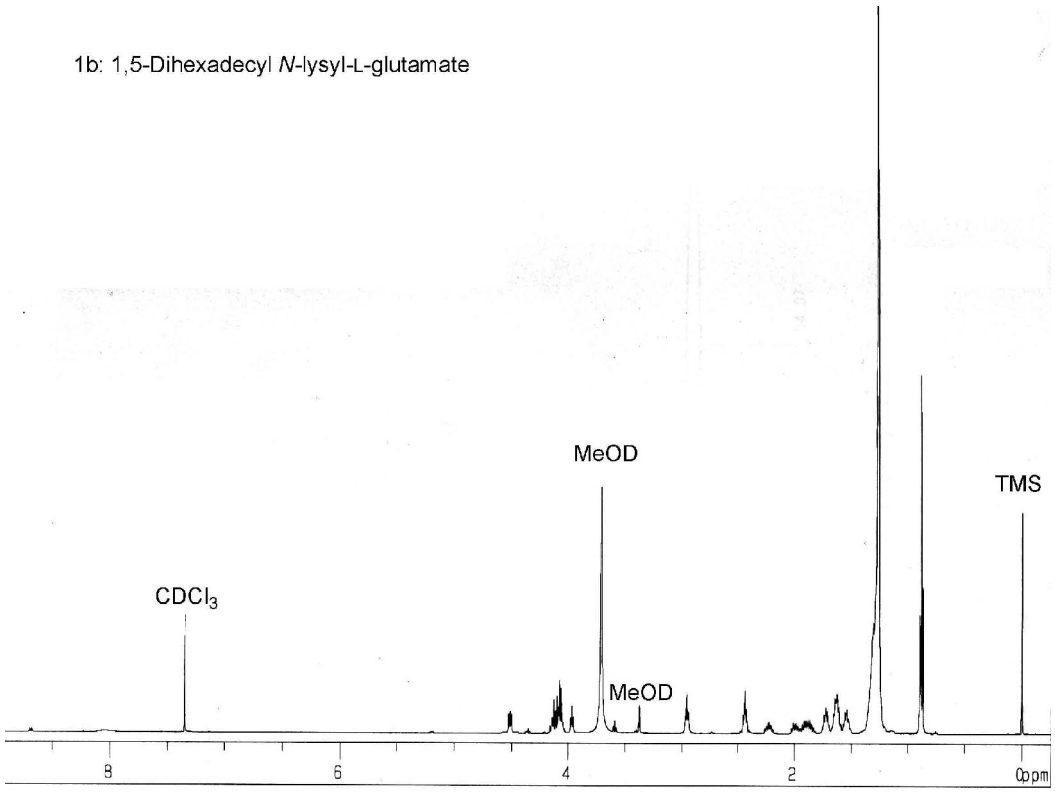

Figure 5S. ^{1}H-NMR spectrum of 1,5-dihexadecyl *N*-lysyl-L-glutamate (**1b**).

1,5-Dihexadecyl *N*-lysyl-L-glutamate (1b): R_f 0.19 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 52H, $CH_{2(palmitoyl)}$), 1.56 (m, 2H, $NH_2CHCH_2CH_2$), 1.62 (m, 4H, $COOCH_2$), 1.69 (m, 2H, $NH_2CH_2CH_2$), 1.87 (m, 2H, NH_2CHCH_2), 1.94-2.25 (m, 2H, $NHCHCH_2$), 2.43 (m, 2H, CH_2COO), 2.95 (m, 2H, NH_2CH_2), 3.97 (m, 1H, NH_2CHCO), 4.02-4.16 (m, 4H, $COOCH_2CH_2$), 4.51 (q, 1H, NHCH). MS(ESI): $(M+H)^+$ calcd. for $C_{43}H_{85}N_3O_5$, 724.15; found, 724.7.

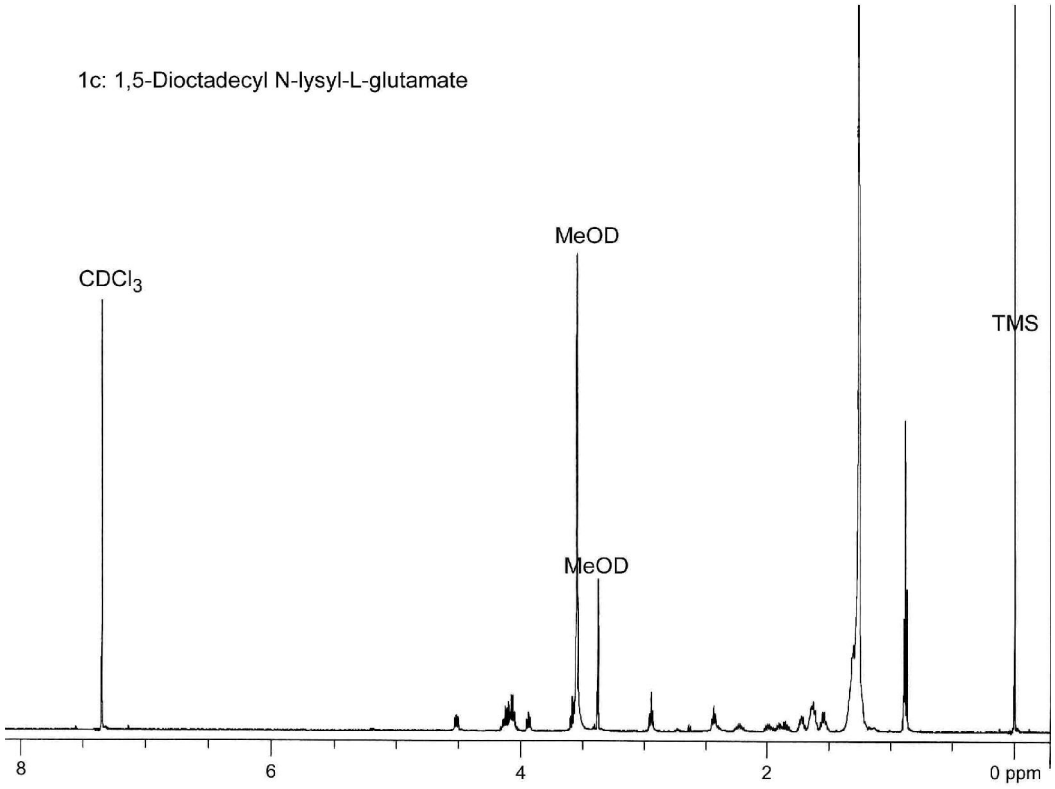

Figure 6S. ^{1}H-NMR spectrum of 1,5-dioctadecyl *N*-lysyl-L-glutamate (**1c**).

1,5-Dioctadecyl *N*-lysyl-L-glutamate (1c): R_f 0.06 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 60H, $CH_{2(stearyl)}$), 1.52 (m, 2H, $NH_2CHCH_2CH_2$), 1.64 (m, 4H, $COOCH_2CH_2$), 1.72 (m, 2H, $NH_2CH_2CH_2$), 1.88 (m, 2H, NH_2CHCH_2), 1.95-2.27 (m, 2H, $NHCHCH_2$), 2.44 (m, 2H, CH_2COO), 2.94 (m, 2H, NH_2CH_2), 3.94 (m, 1H, NH_2CHCO), 4.05-4.12 (m, 4H, $COOCH_2CH_2$), 4.52 (q, 1H, NHCH). MS(ESI): $(M+H)^+$ calcd. for $C_{43}H_{85}N_3O_5$, 780.26; found, 780.7.

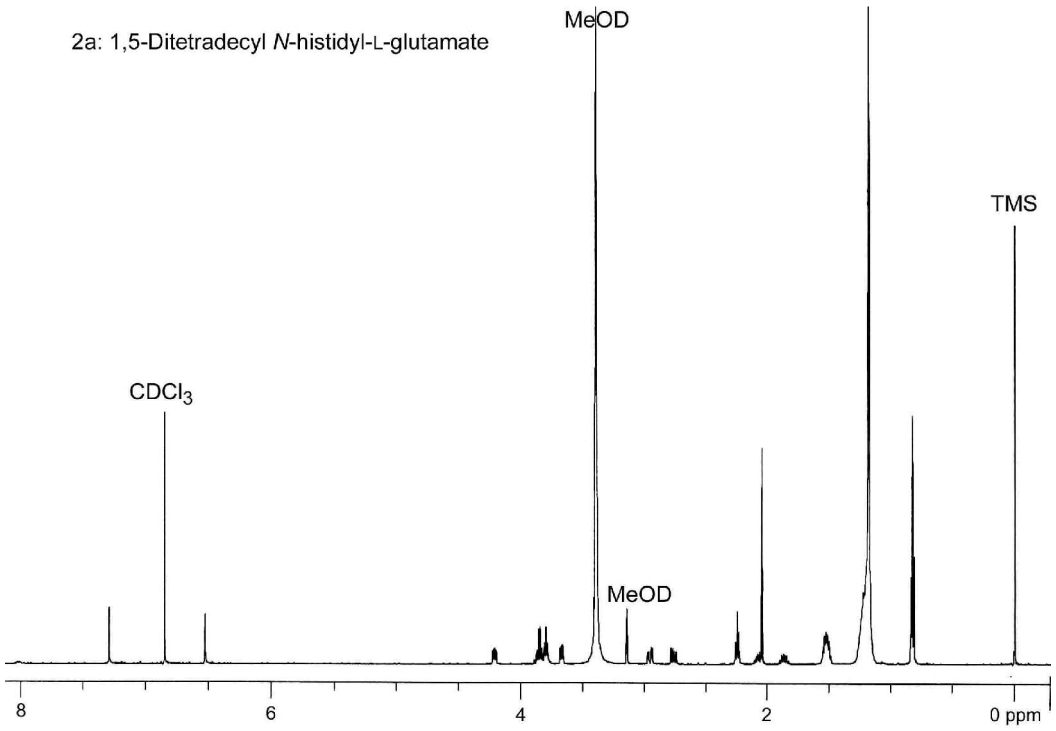

Figure 7S. ^{1}H-NMR spectrum of 1,5-ditetradecyl *N*-histidyl-L-glutamate (**2a**).

1,5-Ditetradecyl *N*-histidyl-L-glutamate (2a): R_f 0.14 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 44H, $CH_{2(myristoyl)}$), 1.64 (m, 4H, $COOCH_2CH_2$), 1.96-2.25 (m, 2H, $NHCHCH_2$), 2.40 (t, 2H, $NHCHCH_2CH_2$), 2.94-3.19 (q, 2H, NH_2CHCH_2), 3.93 (q, 1H, NHCH), 4.05-4.16 (m, 4H, $COOCH_2$), 4.52 (q, 1H, NH_2CH), 7.00 (s, 1H, $CH_2C=CH$), 7.82 (s, 1H, N=CH). MS(ESI): $(M+H)^+$ calcd. for $C_{39}H_{72}N_4O_5$, 677.01; found, 677.6.

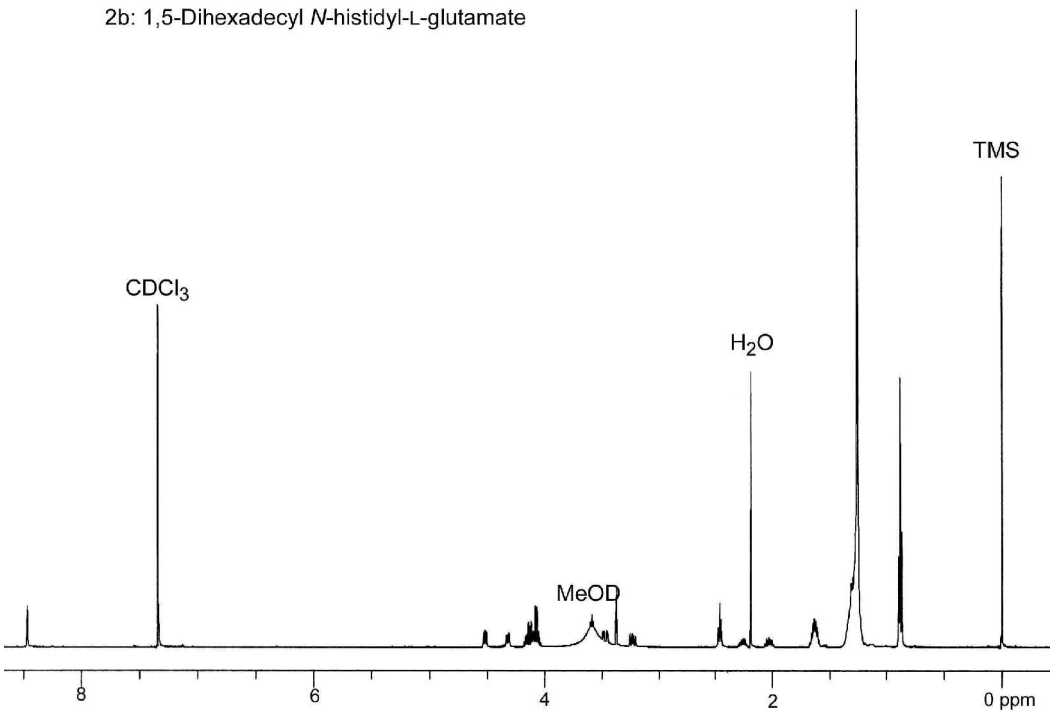

Figure 8S. ^{1}H-NMR spectrum of 1,5-dihexadecyl *N*-histidyl-L-glutamate (**2b**).

1,5-Dihexadecyl *N*-histidyl-L-glutamate (2b): R_f 0.15 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 52H, $CH_{2(palmitoyl)}$), 1.64 (m, 4H, $COOCH_2CH_2$), 1.96-2.29 (m, 2H, $NHCHCH_2$), 2.46 (t, 2H, $NHCHCH_2CH_2$), 3.20-3.49 (q, 2H, NH_2CHCH_2), 4.05-4.14 (m, 4H, $COOCH_2$), 4.33 (q, 1H, NHCH), 4.52 (q, 1H, NH_2CH), 7.82 (s, 1H, N=CH). MS(ESI): $(M+H)^+$ calcd. for $C_{43}H_{80}N_4O_5$, 732.12 ; found, 733.7.

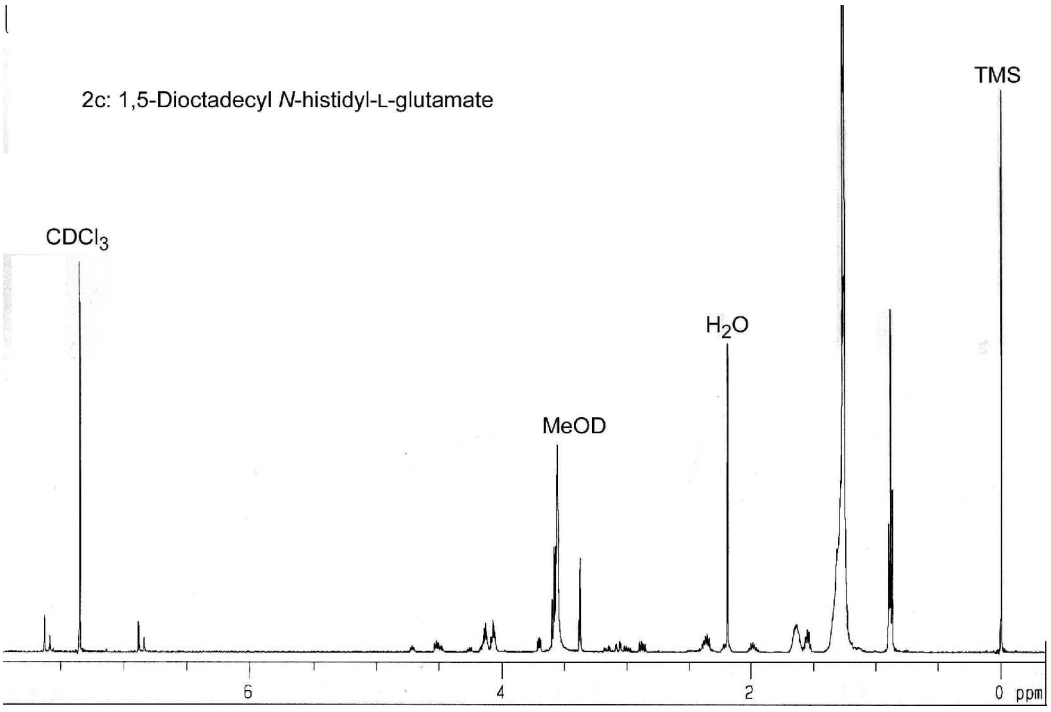

Figure 9S. ^{1}H-NMR spectrum of 1,5-dioctadecyl *N*-histidyl-L-glutamate (**2c**).

1,5-Dioctadecyl *N*-histidyl-L-glutamate (2c): R_f 0.13 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 60H, $CH_{2(stearyl)}$), 1.52-1.63 (m, 4H, $COOCH_2CH_2$), 1.94-2.22 (m, 2H, $NHCHCH_2$), 2.37 (t, 2H, $NHCHCH_2CH_2$), 2.85-3.18 (q, 2H, NH_2CHCH_2), 4.05-4.13 (m, 4H, $COOCH_2$), 4.52 (q, 1H, NH_2CH), 6.88 (d, 1H, $CH_2C=CH$), 7.50 (d, 1H, $N=CH$). MS(ESI): $(M+H)^+$ calcd. for $C_{47}H_{88}N_4O_5$, 789.23; found, 789.7.

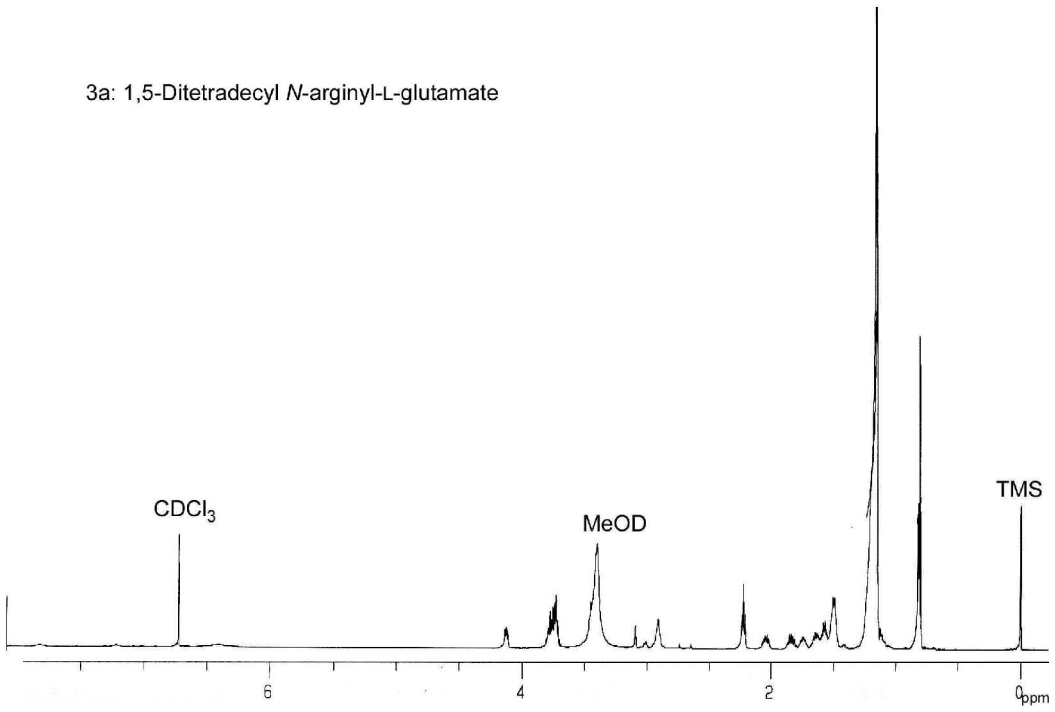

Figure 10S. ^{1}H-NMR spectrum of 1,5-ditetradecyl *N*-arginyl-L-glutamate (**3a**).

1,5-Ditetradecyl *N*-arginyl-L-glutamate (3a): R_f 0.24 ($CHCl_3$:MeOH 4:1). ^{1}H-NMR ($CDCl_3$:MeOD 10:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.27 (m, 44H, $CH_{2(myristoyl)}$), 1.61 (m, 4H, $COOCH_2CH_2$), 1.72 (m, 2H, $NH_2CHCH_2CH_2$), 1.78-1.92 (m, 2H, NH_2CHCH_2), 1.96-2.25 (m, 2H, $NHCHCH_2$), 2.42 (t, 2H, $NHCHCH_2CH_2$), 3.18 (m, 2H, $CH_2NHC{=}N$), 4.06-4.14 (m, 4H, $COOCH_2$), 4.51 (q, 1H, $CONHCH$). MS(ESI): $(M+H)^+$ calcd. for $C_{39}H_{77}N_5O_5$, 696.06; found, 696.7.

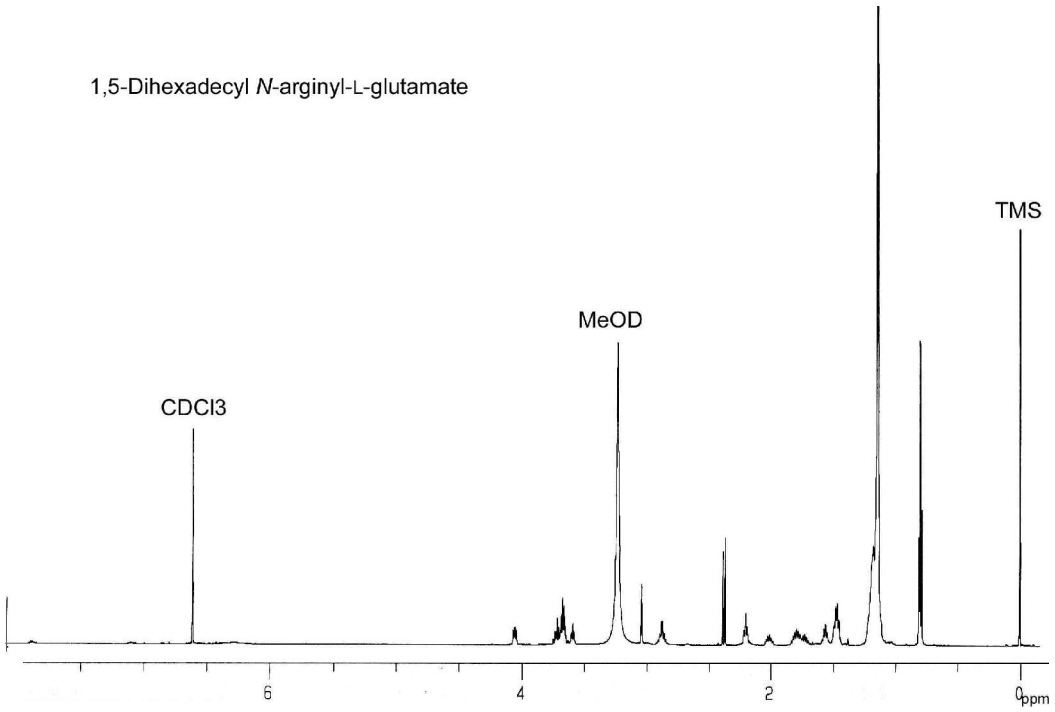

Figure 11S. ^{1}H-NMR spectrum of 1,5-dihexadecyl *N*-arginyl-L-glutamate (**3b**).

1,5-Dihexadecyl *N*-arginyl-L-glutamate (3b): R_f 0.07 ($CHCl_3$:MeOH 10:1). ^{1}H-NMR ($CDCl_3$:MeOD 4:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 52H, $CH_{2(palmitoyl)}$), 1.61 (m, 4H, $COOCH_2CH_2$), 1.72 (m, 2H, $NH_2CHCH_2CH_2$), 1.86-2.02 (m, 2H, $NHCHCH_2$), 1.94-2.26 (m, 2H, $NHCHCH_2$), 2.43 (t, 2H, $NHCHCH_2CH_2$), 3.18 (m, 2H, $CH_2NHC{=}N$), 3.96 (t, 1H, NH_2CHCH_2), 4.06-4.15 (m, 4H, $COOCH_2$), 4.51 (q, 1H, $NHCHCH_2$). MS(ESI): $(M+H)^+$ calcd. for $C_{43}H_{85}N_4O_5$, 752.17; found, 752.8.

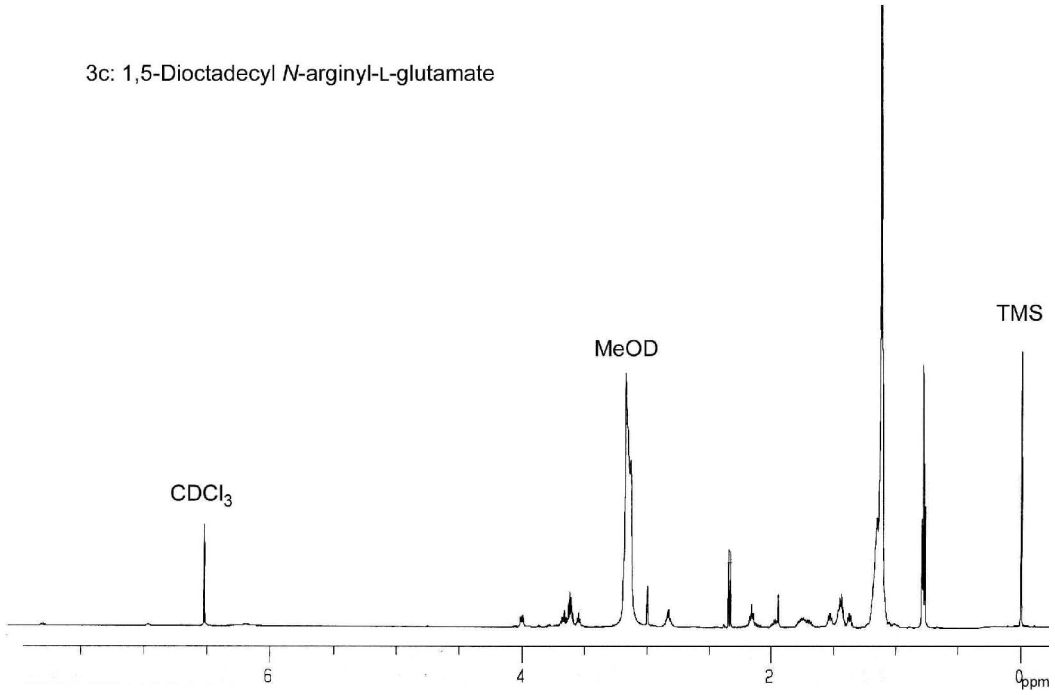

Figure 12S. ^{1}H-NMR spectrum of 1,5-dioctadecyl *N*-arginyl-L-glutamate (**3c**).

1,5-Dioctadecyl *N*-arginyl-L-glutamate (3c): R_f 0.05 ($CHCl_3$:MeOH 101). ^{1}H-NMR ($CDCl_3$:MeOD 4:1, 500 MHz, δ ppm): 0.88 (t, 6H, CH_2CH_3); 1.26 (m, 60H, $CH_{2(stearyl)}$), 1.60 (m, 4H, $COOCH_2CH_2$), 1.52, 1.72 (m, 2H, $NH_2CHCH_2CH_2$), 1.88-2.02 (m, 2H, $NHCHCH_2$), 1.96-2.25 (t, 2H, $NHCHCH_2$), 3.19 (m, 2H, $CH_2NHC{=}N$), 4.00 (t, 1H, $NHCHCH_2$), 4.05-4.16 (m, 4H, $COOCH_2$), 4.51 (q, 1H, NH_2CH). MS(ESI): $(M+H)^+$ calcd. for $C_{39}H_{77}N_5O_5$, 808.27; found, 808.8.