## The Propensity for Secondary Structure of Sialic acid-derived α/δ-Peptides is Sequence and Configuration Specific

Jonel P. Saludes, James B. Ames and Jacquelyn Gervay-Hague\*

Department of Chemistry, University of California, One Shields Ave., Davis, CA 95616

## **Supporting Information**

- S3 Scheme S1: Synthetic scheme for 1
- S3 Scheme S2: Synthetic scheme for 4
- S4 Scheme S3: Representative microwave-assisted solid phase peptide synthesis
- S5-S11 Characterization of compounds 5-11
- S12 Figure S1: Mass spectra of compounds 5-12
- **S14** Figure S2: <sup>1</sup>H NMR spectrum of compound **5**
- **S15** Figure S3: <sup>1</sup>H NMR spectrum of compound **6**
- **S16** Figure S4: <sup>1</sup>H NMR spectrum of compound **7**
- **S17** Figure S5: <sup>1</sup>H NMR spectrum of compound **8**
- **S17** Figure S6: <sup>1</sup>H NMR spectrum of compound **9**
- **S19** Figure S7: <sup>1</sup>H NMR spectrum of compound **10**
- **S20** Figure S8: <sup>1</sup>H NMR spectrum of compound **11**
- S21 Figure S9: <sup>1</sup>H NMR spectrum of compound 12
- **S22** Figure S10: Stack plot of the amide regions from the <sup>1</sup>H NMR DMSO-*d*<sub>6</sub> at 298 K, 600 MHz of **6**, **8**, **10**, **and 12**
- S23 Figure S11: Sample IR spectra and : Plots of amide/azide peak area integration used to monitor SPPS
- **S24** Figure S12: Prominent long range ROE's for **12**
- **S25** Figure S13: Overlay of 25 low energy structures of **8** showing non-convergence to a single-preferred structure

General Experimental Section. NMR spectra were recorded on 400 and 600 MHz NMR spectrometers. Chemical shifts were referenced to residual CHCl<sub>3</sub> ( $\delta_{\rm H}$  = 7.26), CHD<sub>2</sub>OD ( $\delta_{\rm H}$  = 3.30), (CHD<sub>2</sub>)<sub>2</sub>CO ( $\delta_{\rm H}$  = 2.05), CDCl<sub>3</sub> ( $\delta_{\rm C}$  = 77.1), CD<sub>3</sub>OD, ( $\delta_{\rm C}$  = 49.0) or (CD<sub>3</sub>)<sub>2</sub>CO ( $\delta_{\rm C}$  = 39.0). Solutions in D<sub>2</sub>O were referenced to dioxane calibrated to TSP. Low-resolution mass spectra were obtained using a QTRAP mass spectrometer. High Resolution mass spectra were recorded at the UC Davis Molecular Structure Facility using MALDI-TOF with internal calibration. Infrared spectra were recorded on an ATR-FTIR spectrometer. Microwave-assisted SPPS was done using a commercial microwave reactor. Rink amide resin, *N*-acetylneuraminic acid, and Fmoc-Glu(OBut)-OH, and *N*-acetylneuraminic acid were obtained from commercial sources. Reagents were used as received unless otherwise indicated. Solution phase reactions were monitored by TLC using Si gel 60 F254 or RP-C<sub>18</sub> glass-backed plate.

## Synthesis of Sialic Acid Derivative Fmoc-Neu2en (1)

N-(9-fluorenylmethoxycarbonyl)-2,3-dehydro-8,9-isopropylidene neuraminic acid **1** (Fmoc-Neu2en) was synthesized following our previous method<sup>1</sup> with slight modification to furnish an overall yield of 32% in eight steps Scheme 1.



Scheme S1. Synthetic scheme for 1.

## Synthesis of Azidolysine Linker (4)

Compound **4** was prepared following the method developed by Wong,<sup>2</sup> and adopted by others.<sup>3</sup> The cude product was concentrated to dryness, and chromatographed through a column of silica gel using 85:15 CH<sub>2</sub>Cl<sub>2</sub>-MeOH to yield **4** (2.07 g, 92%) as shown in Scheme 2.



Scheme S2. Synthetic scheme for 4.



Scheme S3. Representative microwave-assisted solid phase peptide synthesis.

Assignments of H for each oligomer are based upon the following sytem: **HX-Y**, which refers to **H** in position **X** at residue number **Y** counting from the *N*-terminus.



Neu2en/D-Glu-4 (5): obtained in 27% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): d 5.91 (d, 1H, J = 3.0 Hz, H3-2), 5.82 (d, 1H, J = 3.0 Hz, H3-4), 4.57 (dd, 1H, J = 7.2, 2.4 Hz, H4-2), 4.55 (dd, 1H, J = 7.2, 2.4 Hz, H3-4), 4.54 (dd, 1H, J = 7.2, 4.8 Hz, Hα-3), 4.48 (dd, 1H, J = 10.8, 0.6 Hz, H6-2), 4.39 (dd, 1H, J = 10.8, 0.6 Hz, H6-4), 4.24 (dd, 1H, J = 10.8, 9.0 Hz, H5-2), 4.16 (dd, 1H, J = 10.8, 8.4 Hz, H5-2), 4.15 (dd, 1H, J = 4.8, 2.4 Hz, Hα-5), 4.14 (dd, 1H, J = 4.8, 2.4 Hz, Hα-2), 4.00 (ddd, 1H, J = 9.6, 5.4, 2.4 Hz, H8-2), 3.92-3.88 (m, 3H, H8-4, H9/9'-2), 3.71 (dd, 1H, J = 6.0, 6.0 Hz, H9'-4), 3.69 (dd, 1H, J = 3.0, 3.0 Hz, H7-2), 3.67 (dd, 1H, J = 6.0, 6.0 Hz, H9-4), 3.69 (dd, 1H, J = 3.0, 3.0 Hz, H7-2), 2.62-2.49 (m, 4H, Hγ-1/3), 2.28-2.09 (m, 4H, Hβ-1/3), 1.90-1.78 (m, 2H, Hβ-5), 1.66-1.58 (m, 2H, Hδ-5), 1.44 (p, 2H, J = 7.8 Hz, Hγ-5). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): d 180.2, 179.3, 178.3, 176.2, 172.8, 166.4, 148.7, 148.0, 111.7, 110.9, 78.8, 72.7, 70.7, 70.2, 69.8, 65.8, 65.7, 56.5, 55.9, 55.8, 52.8 42.0, 33.6, 33.1, 32.6, 30.6, 29.1, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>34</sub>H<sub>53</sub>N<sub>9</sub>O<sub>19</sub>Na<sup>+</sup>, 914.3350; found, 914.3400.



Neu2en/D-Glu-6 (6): obtained in 39% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): d 5.91 (d, 1H, J = 2.4 Hz, H3-2), 5.89 (d, 1H, J = 2.4 Hz, H3-4), 5.82 (d, 1H, J = 1.8 Hz, H3-6), 4.57-4.51 (m, 5H, H4-6, Hα-5, H4-4, Hα-3, H4-2), 4.48 (dd, 1H, J = 10.8, 1.2 Hz, H6-2), 4.43 (dd, 1H, J = 10.8, 1.2 Hz, H6-4), 4.38 (dd, 1H, J = 10.8, 1.2 Hz, H6-6), 4.24 (dd, 1H, J = 10.8, 9.0 Hz, H5-2), 4.20-4.18 (m, 1H, H5-4), 4.15 (dd, 1H, J = 9.0, 1.8 Hz, H5-6), 4.14 (t, 1H, J = 5.4 Hz, Hα-7), 4.14-4.12 (m, 1H, Hα-1), 4.01-3.96 (m, 3H, H8-2/4), 3.91-3.85 (m, 5H, H8-6, H9-2/4), 3.72-3.61 (m, 5H, H7-6, H9-6, H7-4, H7-2), 3.37-3.28 (m, 2H, Hε-7), 2.58-2.47 (m, 6H, Hγ-1/3/5), 2.27-2.09 (m, 6H, Hβ-1/3/5), 1.88-1.78 (m, 2H, Hβ-7), 1.61 (p, 2H, J = 6.6 Hz, Hδ-7), 1.44 (p, 2H, J = 7.8 Hz, Hγ-7). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): d 180.7, 180.4, 179.9, 178.4, 176.5, 172.7, 166.6, 166.0, 148.6, 148.0, 111.9, 111.7, 110.9, 79.0, 78.9, 72.9, 72.8, 70.8, 70.7, 69.7, 69.6, 65.9, 65.8, 65.7, 56.6, 55.8, 52.9, 52.8, 41.9, 33.8, 33.4, 33.0, 30.7, 29.4, 29.3, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>48</sub>H<sub>73</sub>N<sub>11</sub>O<sub>28</sub>Na<sup>+</sup>, 1274.4518; found, 1274.452.



**D-Glu/Neu2en-4 (7):** obtained in 39% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): d 5.97 (d, 1H, J = 2.4 Hz, H3-1), 5.90 (d, 1H, J = 2.4 Hz, H3-3), 4.66 (dd, 1H, J = 8.4, 3.0 Hz, H4-1), 4.65 (dd, 1H, J = 10.2, 1.8 Hz, H6-1), 4.58 (dd, 2H, J = 9.0, 2.4 Hz, H4-3, Hα-2), 4.45 (dd, 1H, J = 10.8, 0.6 Hz, H6-3), 4.43 (dd, 1H, J = 9.0, 5.4 Hz, Hα-4), 4.20 (dd, 1H, J = 10.8, 9.0 Hz, H5-3), 4.15 (dd, 1H, J = 7.2, 5.4 Hz, Hα-5), 4.08 (ddd, 1H, J = 9.0, 5.4, 2.4 Hz, H8-1), 4.00 (ddd, 1H, J = 9.6, 6.0, 2.4 Hz, H8-3), 3.94 (dd, 1H, J = 12.0, 2.4 Hz, H9'-1), 3.92 (dd, 1H, J = 10.8, 2.4 Hz, H7-1), 3.91 (dd, 1H, J = 9.0, 1.8 Hz, H9-3), 3.71 (dd, 1H, J = 12.0, 6.0 Hz, H9'-3), 3.66 (dd, 1H, J = 9.6, 1.2 Hz, H7-3), 3.65 (dd, 1H, J = 10.2, 8.4 Hz, H5-1), 3.31-3.22 (m, 2H, Hε-5), 2.60-2.48 (m, 4H, Hγ-2/4), 2.30-2.06 (m, 4H, Hβ-2/4), 1.91-1.77 (m, 2H, Hβ-5), 1.64-1.52 (m, 2H, Hδ-5), 1.44 (p, 2H, J = 7.2 Hz, Hγ-5). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): d 180.3, 180.2, 178.4, 176.3, 175.7, 166.4, 166.0, 145.0, 111.8, 111.0, 79.0, 77.9, 72.9, 72.8, 70.8, 70.6, 69.7, 67.6, 66.0, 65.8, 65.5, 56.6, 56.4, 53.7, 52.9, 41.8, 33.8, 33.1, 30.8, 29.2, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>34</sub>H<sub>53</sub>N<sub>9</sub>O<sub>19</sub>Na<sup>+</sup>, 914.3350; found, 914.3410.



**D-Glu/Neu2en-6 (8):** obtained in 24% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz):  $\delta$  5.98 (d, 1H, J = 2.4 Hz, H3-1), 5.92 (d, 1H, J = 2.4 Hz, H3-3), 5.90 (d, 1H, J = 2.4 Hz, H3-5), 4.67 (dd, 1H, J= 6.6, 2.4 Hz, H4-1), 4.66 (dd, 1H, J = 7.8, 2.4 Hz, H6-1), 4.59 (dd, 2H, J = 9.0, 2.4 Hz, H4-3/5), 4.58 (dd, 1H, J = 9.0, 4.8 Hz, H $\alpha$ -2), 4.56 (dd, 1H, J = 8.4, 4.2 Hz, H $\alpha$ -4), 4.47 (dd, 1H, J = 7.2, 1.2 Hz, H6-3), 4.45 (dd, 1H, J = 7.2, 1.2 Hz, H6-5), 4.42 (dd, 1H, J = 9.0, 5.4 Hz, H $\alpha$ -6), 4.22 (dd, 1H, J = 9.0, 4.8 Hz, H5-3), 4.20 (dd, 1H, J = 8.4, 4.2 Hz, H5-5), 4.16 (dd, 1H, J = 7.8, 5.4 Hz, H $\alpha$ -7), 4.07 (ddd, 1H, J = 9.0, 5.4, 3.0 Hz, H8-1), 4.03-3.99 (m, 2H, H8-3/5), 3.95 (dd, 1H, J = 12.0, 3.0 Hz, H9'-1), 3.93 (dd, 2H, J = 11.4, 2.4 Hz, H9'-3/5), ), 3.92 (dd, 1H, J = 8.4, 1.8Hz, H7-1), ), 3.80 (dd, 1H, J = 12.6, 6.0 Hz, H9-1), 3.73 (dd, 1H, J = 6.6, 1.8 Hz, H9-3), 3.71 (dd, 1H, J = 7.8, 1.8 Hz, H9-5), 3.67 (dd, 1H, J = 9.0, 1.2 Hz, H7-3), 3.66 (dd, 1H, J = 9.0, 1.2 Hz, H5-5), 3.32-3.22 (m, 2H, Hε-7), 2.60-2.49 (m, 6H, Hγ-2/4/6), 2.31-2.07 (m, 6H, Hβ-2/4/6), 1.91-1.78 (m, 2H, Hβ-7), 1.63-1.56 (m, 2H, Hδ-7), 1.44 (p, 2H, J = 7.8 Hz, Hγ-7). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): δ 180.3, 178.3, 176.2, 175.6, 166.2, 166.1, 165.8, 148.1, 148.0, 111.8, 110.9, 79.0, 77.8, 72.9, 72.6, 70.7, 70.5, 69.6, 67.5, 65.9, 65.8, 65.4, 56.5, 53.7, 52.7, 41.8, 33.7, 33.3, 30.7, 29.3, 24.8. MALDI-TOFMS:  $[M + Na^+]$  calcd for  $C_{48}H_{73}N_{11}O_{28}Na^+$ , 1274.4518; found, 1274.452.



Neu2en/L-Glu-4 (9): obtained in 28% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): d 5.89 (d, 1H, J = 2.4 Hz, H3-2), 5.83 (d, 1H, J = 1.8 Hz, H3-4), 4.53 (dd, 1H, J = 9.0, 2.4 Hz, H4-2), 4.52 (dd, 1H, J = 9.0, 2.4 Hz, H4-4), 4.49 (dd, 1H, J = 9.0, 5.4 Hz, Hα-1), 4.42 (d, 1H, J = 10.8 Hz, H6-2), 4.39 (d, 1H, J = 10.8 Hz, H6-4), 4.22 (dd, 1H, J = 10.8, 9.0 Hz, H5-2), 4.16-4.12 (m, 2H, Hα-5, H5-4), 4.11 (m, 1H, Hα-3), 4.00 (ddd, 1H, J = 9.6, 5.4, 2.4 Hz, H8-2), 3.96-3.86 (m, 2H, H8-4, H9'-2), 3.72 (d, 1H, J = 10.2 Hz, H7-2), 3.70-3.69 (m, 2H, H9-2/4), 3.68 (d, 1H, J = 10.8 Hz, H7-4), 3.38-3.28 (m, 2H, Hε-5), 2.66-2.51 (m, 4H, Hγ-1/3), 2.35-2.10 (m, 4H, Hβ-1/3), 1.89-1.77 (m, 2H, Hβ-5), 1.65-1.57 (m, 2H, Hδ-5), 1.44 (p, 2H, J = 7.8 Hz, Hγ-5). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): d 180.0, 179.0, 178.1, 176.2, 172.4, 166.3, 166.1, 148.2, 147.6, 111.8, 110.8, 78.9, 72.7, 72.5, 70.6, 70.5, 69.7, 65.8, 65.7, 56.4, 55.6, 52.9, 41.9, 33.7, 32.8, 32.0, 30.6, 28.9, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>34</sub>H<sub>53</sub>N<sub>9</sub>O<sub>19</sub>Na<sup>+</sup>, 914.3350; found, 914.3320.



**Neu2en/L-Glu-6 (10):** obtained in 15% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): δ 5.88 (d, 1H, J = 2.4 Hz, H3-2), 5.87 (d, 1H, J = 2.4 Hz, H3-4), 5.82 (d, 1H, J = 2.4 Hz, H3-6), 4.54 (dd, 3H, 9.0, 3.0 Hz, H4-2/4/6), 4.43-4.37 (m, 5H, H6-6, Hα-5, H6-4, Hα-3, H6-2), 4.22 (dd, 1H, J = 10.2, 8.4 Hz, H5-2), 4.17-4.11 (m, 3H, Hα-7, H5-4/6), 4.08 (t, 1H, J = 6.6 Hz, Hα-1), 4.01 (ddd, 1H, J = 9.0, 5.4, 2.4 Hz, H8-2), 3.97 (ddd, 1H, J = 12.0, 6.0, 2.4 Hz, H8-4), 3.92-3.86 (m, 3H, H8-C<sub>a</sub>, H9'-2/4), 3.91-3.85 (m, 7H, H7-2/4/6, H9-2/4/6, H9'-6), 3.36-3.28 (m, 2H, Hε-7), 2.45-2.35 (m, 6H, Hγ-1/3/5), 2.18-2.07 (m, 6H, Hβ-1/3/5), 1.78-1.88 (m, 2H, Hβ-7), 1.65-159 (m, 2H, Hδ-7), 1.44 (p, 2H, J = 7.8 Hz, Hγ-7). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): δ 179.9, 179.8, 178.7, 177.9, 176.0, 172.4, 166.2, 166.1, 166.0, 148.1, 147.4, 111.7, 111.5, 111.0, 78.8, 72.7, 72.5, 70.5, 69.5. 66.5, 65.8, 65.6, 57.1, 56.0, 52.8, 52.7, 41.8, 35.6, 33.7, 30.7, 30.1, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>48</sub>H<sub>73</sub>N<sub>11</sub>O<sub>28</sub>Na<sup>+</sup>, 1274.4518; found, 1274.456.



**L-Glu/Neu2en-4 (11):** obtained in 24% purified yield. <sup>1</sup>H NMR (D<sub>2</sub>O, 600 MHz): d 5.96 (d, 1H, J = 2.4 Hz, H3-1), 5.91 (d, 1H, J = 2.4 Hz, H3-3), 4.67 (dd, 1H, J = 8.4, 3.0 Hz, H4-1), 4.64 (dd, 1H, J = 9.6, 2.4 Hz, H6-1), 4.57 (dd, 1H, J = 9.0, 2.4 Hz, H4-3), 4.52 (dd, 1H, J = 8.4, 5.4 Hz, Hα-2), 4.44 (dd, 1H, J = 7.2, 4.2 Hz, Hα-4), 4.44 (dd, 1H, J = 10.8, 1.2 Hz, H6-3), 4.19 (dd, 1H, J = 10.8, 9.0 Hz, H5-3), 4.15 (dd, 1H, J = 7.2, 5.4 Hz, Hα-5), 4.08 (ddd, 1H, J = 9.0, 6.0, 3.0 Hz, H8-1), 4.00 (ddd, 1H, J = 9.0, 6.0, 3.0 Hz, H8-3), 3.94 (dd, 1H, J = 12.0, 3.0 Hz, H9'-3), 3.92 (dd, 1H, J = 8.4, 1.2 Hz, H7-1), 3.91 (dd, 1H, J = 12.0, 3.0 Hz, H9'-3), 3.80 (dd, 1H, J = 12.0, 5.4 Hz, H9-1), 3.72 (dd, 1H, J = 9.6, 1.2 Hz, H7-3), 3.71 (dd, 1H, J = 12.0, 6.0 Hz, H9-3), 3.66 (dd, 1H, J = 7.2, 7.2 Hz, Hγ-2/4), 2.28-2.07 (m, 4H, Hβ-2/4), 1.90-1.78 (m, 2H, Hβ-5), 1.64-1.55 (m, 2H, Hδ-5), 1.44 (p, 2H, J = 7.8 Hz, Hγ-5). <sup>13</sup>C NMR (D<sub>2</sub>O, 150 MHz): d 180.3, 180.2, 178.4, 176.5, 175.6, 166.3, 165.9, 148.1, 148.0, 112, 110.9, 79.1, 77.9, 72.8, 72.7, 70.7, 70.5, 69.7, 67.6, 66.0, 65.8, 65.4, 56.4, 56.3, 53.7, 52.8, 41.9, 33.7, 33.0, 30.6, 29.2, 29.0, 24.8. MALDI-TOFMS: [M + Na<sup>+</sup>] calcd for C<sub>34</sub>H<sub>53</sub>N<sub>9</sub>O<sub>19</sub>Na<sup>+</sup>, 914.3350; found, 914.3330.

Figure S1. Mass spectra of  $\alpha/\delta$ -hybrid peptides (5-12).



 $\begin{array}{l} C_{34}H_{53}N_9O_{19}Na^+ \\ Found = 914.3410 \\ Theor. = 914.3350 \end{array}$ 























**Figure S10**. Stack plot of the amide regions from <sup>1</sup>H NMR spectra (DMSO- $d_6$  at 298 K, 600 MHz) showing that the amide H of **12** are more dispersed and distinct from each other compared to **6**, **8**, and **10**.



**Figure S11.** (A) Stack plot of IR spectra for on-bead monitoring of SPPS: absorption bands at 1719 cm<sup>-1</sup> due to C=O of Fmoc, 2101 cm<sup>-1</sup> due to N<sub>3</sub>. (B) and (C) Plots of amide/azide IR band peak area integration showing the linear correlation as a function of peptide length: (B) Odd numbered\* residues, (C) even numbered\* residues. red plot: oligomer with Fmoc protected *N*-terminus; blue plot: oligomer with deprotected *N*-terminus.

\**N.B.* Number system is based on the growing peptide chain bound to  $N_3Lys$  on the bead, *i.e.* Glu6 is loaded residue number 1.



**Figure S12.** Crucial long range nOe's for **12** extracted from the ROESY spectra (9:1 H<sub>2</sub>O/D<sub>2</sub>O), Neu2en-H8(*i*) $\rightarrow$ L-Glu HN(*i*+1) nOe's are in blue, Neu2en-H3(*i*) $\rightarrow$ Neu2en-H9'(*i*+1) are in brown, L-Glu2-H $\alpha$  $\rightarrow$  L-Glu4-HN is in green, and N<sub>3</sub>Lys7-H2 $\rightarrow$ L-Glu4-H $\gamma$  is in black.

Figure S12: Prominent long range ROE's for **12**.

**Figure S13.** Overlay of 25 low energy structures of **8** showing non-convergence to a single-preferred structure; all H have been omitted for clarity (backbone atom rmsd:  $2.080 \pm 0.750$ ; non-hydrogen atom rmsd:  $3.325 \pm 1.013$ ).

References:

- (1) Gregar, T. Q.; Gervay-Hague, J. J. Org. Chem. 2004, 69, 1001-1009.
- (2) Nyffeler, P. T.; Liang, C.-H.; Koeller, K. M.; Wong, C.-H. J. Am. Chem. Soc. 2002, 124, 10773-10778.
- (3) Lundquist, J. T.; Pelletier, J. C. Org. Lett. 2001, 3, 781-783.