Supporting Information for:
Controlling the transmembrane transport of nucleosides

Gretchen Marie Peters and Jeffery T. Davis*

Department of Chemistry and Biochemistry, University of Maryland

College Park, MD 20742, USA

*jdavis@umd.edu

Table of Contents

General Experimental	S2
Synthetic Procedures	S2-3
Fig. S1. ¹ H NMR spectrum of G 1	S4
Fig. S2. ¹³ C NMR spectrum of G 1	S5
Fig. S3. ESI-MS of G 1	S6
Fig. S4. ¹ H NMR spectrum of C 1	S7
Fig. S5. ¹³ C NMR spectrum of C 2	S8
Fig. S6. ESI-MS of C 2	S9
Experimental Procedure for Phase Transfer Experiments	S10
Fig. S7. ¹ H-NMR spectra of G 1 and (G 1) ₈ •K ⁺ (2,6-DNP) ⁻ in CDCl ₃	S11
Fig. S8. ¹ H NMR titrations of C 2 with G 1 in CDCl ₃	S12
Fig. S9. ESI-MS of G 1 •C 2 basepair	S13
Fig. S10. Inhibition of dG 7 transport by formation of (G 1)•C 2 •G 1 triple	S14
Fig. S11. Dependence of rC 3 and araC 5 transport on the Na ₂ B ₄ O ₇ concentration	S15
Fig. S12. Reinitiation of rC 3 transport with the addition of ribose 10	S16
Fig.S13. DOSY spectra for Borate esters 11-13	S17
Table 1. ¹ H NMR chemical shifts for Borate Esters 11-13	S18
Table 2. ¹ H NMR Diffusion Coefficients for Borates 11-13	S19
Supporting Information References	S19

General experimental: All ^1H NMR spectra were recorded on a Bruker DRX-400 operating at 400.13 MHz or a Bruker AVIII-600 operating at 600.13 MHz. All ^{13}C NMR spectra were recorded on a Bruker DRX-500 operating at 500.13 MHz. The ^1H and ^{13}C chemical shifts are reported in ppm relative to the residual solvent peak. ^{11}B NMR were recorded on a Bruker DRX-500 operating at 500.13 MHz or a Bruker AVIII-600 operating at 600.13 MHz. ^{11}B chemical shifts are reported in ppm relative to $\text{BF}_3\cdot\text{O}(\text{C}_2\text{H}_5)_2$. ESI-MS experiments were done with a JEOL AccuTOF spectrometer. Chromatography was performed using 60-200 mesh silica gel from Baker. Thin layer chromatography was performed on Kieselgel 60 F254 silica-coated glass plates and visualized by UV-vis lamp. Deuterated solvents were purchased from Cambridge Isotope Labs. Chemicals and solvents were purchased from Santa Cruz Biotechnology, Aldrich, Fisher, and Acros.

Synthetic Procedures:

Synthesis of 2',3',5'-O-trihexanoyl guanosine G 1. In a clean, dry round bottom flask, guanosine (1.241 g, 4.381 mmol) was suspended in dry acetonitrile (15 mL). DMAP (0.0155 g, 0.1269 mmol), triethylamine (2.44 mL, 0.01749 mol) and hexanoic anhydride (3.65 mL, 0.01577 mol) were then added. The reaction mixture was stirred under N_2 until the solution turned clear, in approximately 30 minutes. The reaction progress was monitored by TLC using 92:8 CH_2Cl_2 :MeOH ($R_f=0.27$). Upon quenching of the reaction with water, a white precipitate formed. The solution was vacuum filtered, the precipitate was washed with water and then dried to give G 1 as a white powder (2.31 g, 91.4 %). ^1H NMR (400.13 MHz, d_6 -DMSO, ppm) δ 10.75 (s, 1 H), 7.90 (s, 1 H), 6.53 (s, 2 H), 5.97 (d, 1 H, $J=5.8$ Hz), 5.81 (t, 1 H, $J=6.0$ Hz), 5.51 (m, 1 H), 4.32 (m, 3 H), 2.32 (m, 6 H), 1.52 (m, 6 H), 1.27 (m, 12 H), 0.83 (m, 9 H). ^{13}C NMR (500.13 MHz, CDCl_3 , ppm) δ 173.61, 172.57, 172.31, 159.24, 154.11, 151.67, 136.50, 117.76, 86.41, 80.44, 72.93, 70.78, 63.24, 34.17, 34.06, 33.89, 31.41, 24.65, 24.53, 22.46, 22.44, 22.41, 14.03, 14.00. MS (ESI $^+$): $\text{C}_{28}\text{H}_{43}\text{N}_5\text{O}_8$, Mass Calculated: 577.31, Mass Found (M+1): 578.31.

Synthesis of 2',3',5'-O-trihexanoyl cytidine C 2. In a clean, dry round bottom flask, cytidine (1.23 g, 4.34 mmol) was suspended in dry acetonitrile (15 mL). DMAP (0.0080 g, 0.065 mmol), triethylamine (2.42 mL, 0.0174 mol) and hexanoic anhydride (3.62 mL, 0.0156 mol) were then added. The reaction mixture was stirred under N₂ until the solution turned clear, approximately in 3 hours. The reaction progress was monitored by TLC using 95:5 CH₂Cl₂:MeOH (R_f=0.57). The solution was then suspended in a 1:1 mixture of CH₂Cl₂:H₂O. The organic layer was extracted and washed three times with water. Column chromatography was performed with increasing increments of 0-5%MeOH:CH₂Cl₂. Fractions were combined and dried *in vacuo* to give **2** as a white powder (0.404 g, 49.9 %). ¹H NMR (400.13 MHz, CDCl₃, ppm) 7.43 (d, 1 H, J= 7.6 Hz), 6.00 (d, 1 H, J= 4.3 Hz), 5.90 (d, 1 H, J= 7.4 Hz), 5.39 (m, 2 H), 4.34 (m, 3 H) 2.34 (m, 6 H), 1.63 (m, 6 H), 1.30 (m, 12 H), 0.89 (m, 9 H). ¹³C NMR (500.13 MHz, CDCl₃) δ 173.41, 172.71, 172.53, 166.23, 155.72, 141.20, 96.26, 90.05, 79.58, 73.42, 70.22, 63.29, 34.23, 34.04, 33.99, 31.42, 31.40, 31.37, 24.67, 24.59, 24.57, 22.45, 22.43, 14.02. MS (ESI⁺): C₂₇H₄₃N₃O₈, Mass Calculated: 537.31, Mass Found (M+1): 538.22.

Preparation of G-Quartet octamer (G 1)₈•K⁺DNP⁻ from 2',3',5'-O-trihexanoyl guanosine: In a clean, dry round bottom flask, 2',3',5'-O-trihexanoyl guanosine G **1** (0.234 g, 0.415 mmol) was dissolved in CH₂Cl₂ (25 mL). Potassium 2,6-dinitrophenol (0.047 g, 0.212 mmol) was dissolved in H₂O (25 mL) and the aqueous solution was added to the flask. The resulting biphasic solution was stirred under N₂ gas for ~3 hours. The aqueous layer was removed and discarded. The remaining organic layer was concentrated and dried to give a translucent, orange gel. ¹H NMR (600.13 MHz, CDCl₃) δ 12.12 (s, 4 H), 12.01 (s, 4 H), 8.24 (d, 2 H, DNP), 7.75 (s, 4 H), 7.57 (s, 4 H), 6.76 (t, 1 H, DNP), 6.14 (t, 4 H, J= 5.1 Hz), 6.11 (d, 4 H, J= 4.8 Hz), 5.97 (dd, 4 H J= 6.7, 1.9 Hz), 5.93 (d, 4H, J= 1.9 Hz), 5.76 (t, 4 H, J= 5.4 Hz), 5.60 (dd, 4 H, J= 8.4, 6.9 Hz), 4.58 (m, 4 H), 4.48 (m, 8 H), 4.40 (m, 4 H), 4.34 (m, 4 H), 2.40 (m, 48 H), 1.65 (m, 48 H), 1.32 (m, 96 H), 0.92 (m, 72 H).

Figure S1. ^1H NMR (400.13 MHz) spectrum of 2', 3', 5'-O-trihexanoyl guanosine **G 1** in DMSO-d_6 .

Figure S2. ^{13}C NMR (500.13 MHz) spectrum of 2', 3', 5'-O-trihexanoyl guanosine **G 1** in CDCl_3 .

Figure S3. ESI-MS (positive mode) of 2', 3', 5'-O-trihexanoyl guanosine G **1**. Calculated mass for $C_{28}H_{43}N_5O_8$: 577.31; found $(M+H)^+$: 578.31.

Figure S4. ^1H NMR (400.13 MHz) spectrum of 2', 3', 5'-O-trihexanoyl cytidine C 2 in CDCl_3 .

Figure S5. ^{13}C NMR (500.13 MHz) spectrum of 2', 3', 5'-O-trihexanoyl cytidine C 2 in CDCl_3 .

Figure S6. ESI-MS (positive mode) of 2', 3', 5'-O-trihexanoyl cytidine C 2. Calculated mass for $C_{27}H_{43}N_3O_8$: 537.31; found $(M+H)^+$: 538.22.

Experimental Procedure of Phase Transfer Experiments:

A solution of carrier (20 mM) in chloroform (3 mL) was placed in the bottom of a U-tube apparatus (8 mm internal diameter, 7.5 cm high, 2.5 cm between each arm) equipped with a stir bar on a magnetic stirrer. A solution of the target nucleoside(s) (3-25 mM) was prepared in H₂O, D₂O or aqueous buffer and was carefully added to the source arm (1.5 mL) of the U-tube. In the opposite, receiving arm, D₂O (1.5 mL) was added. The U-tube was clamped above a magnetic stirrer and the lower CHCl₃ layer was magnetically stirred at room temperature, making sure that no mechanical transfer of material occurred between the source and receiving phases. The transport of nucleosides across the liquid membrane was monitored by ¹H NMR spectroscopy of the receiving phase, which allowed us to distinguish between nucleosides (e.g. a doublet at $\delta = 5.91$ ppm for H1' of rC **3** and a triplet at $\delta = 6.26$ ppm for H1' of dC **4**) and measure the relative amounts of transported nucleosides. Every 3 hours, a 0.5 mL aliquot was removed from the receiving arm, and its ¹H NMR spectrum was recorded. The concentration of the transported compounds in the receiving phase was determined relative to an internal standard of a known concentration of adenosine that was contained within a capillary tube. After NMR analysis, the sample was returned to the U tube. All transport runs were reproduced at least in duplicate and usually in triplicate.

Figure S7. In the presence of $\text{K}^+ \text{DNP}^-$ **8**, lipophilic **G 1** forms an octamer, $(\text{G } \mathbf{1})_8 \cdot \text{K}^+ \text{DNP}^-$. The ^1H NMR spectrum of monomeric lipophilic **G 1** in CDCl_3 is shown above (top). Upon the addition of $\text{K}^+ \text{DNP}^-$ **8** to a solution of lipophilic **G 1**, the protons are split into two sets of signals in the ^1H NMR spectrum (bottom). This NMR pattern is a classic signature of G-quartet formation. [1] The structure of the octamer $(\text{G } \mathbf{1})_8 \cdot \text{K}^+ \text{DNP}^-$ was further verified by 2D NOESY.

Figure S8. ^1H NMR titrations of G **1** with a constant concentration of C **2** (10 mM) in CDCl_3 . The most significant shifts are seen in the exchangeable protons. Initially, the G **1** imino NH1 (black) undergoes an upfield shift. With an excess of G **1**, the imino peak broadens and begins to shift back downfield. The exchangeable amino protons for G **1** and C **2** are represented in red, blue, and green and undergo significant shifts and sharpening. Interestingly, there is also an observable shift in the sugar C **2** H1' signal (orange). The nucleobase peaks of both lipophilic nucleosides also undergo shifts. The H8 of G **1** shifts downfield and both the H5 and H6 of C **2** shift upfield. These results are consistent with the formation of a stable lipophilic base pair G **1**•C **2** and base triple G **1**•C **2**•G **1**. [2]

Figure S9. ESI-MS (positive mode) of a 1:1 mixture of 2', 3', 5'-O-trihexanoyl guanosine G 1 and 2', 3', 5'-O-trihexanoyl cytidine C 2. Calculated mass for G 1•C 2 1114.62; found (G 1•C 2 + H)⁺: 1115.54.

Figure S10. The $G 1 \cdot C 2$ basepair transports dG 7 (triangles), but dG 7 is not transported by a complex with formula $(G 1)_2 \cdot C 2$ (circles). In these experiments, the source phase contained a solution of dG 7 (3.25 mM), and the $CHCl_3$ phase contained either a) an equimolar amount of lipophilic G 1 (20 mM) and lipophilic C 2 (20 mM) or b) a 2:1 ratio of lipophilic G 1 (40 mM) to lipophilic C 2 (20 mM). To ensure that these results were not the result of transport of dG 7 by any “free” C 2, dC 3 (3.25 mM) was also present in the source phase as a control. We observed no transport of dC 3 by the $G 1 \cdot C 2$ basepair. Transport of dG 7 by the lipophilic basepair $G 1 \cdot C 2$ (triangles) was inhibited when the ratio of G 1 to C 2 was changed from 1:1 to 2:1 (circles). We attribute this inhibition to the formation of a $G 1 \cdot C 2 \cdot G 1$ triple in the organic phase, which precludes binding and transport of dG 7.

Figure S11. Dependence of the carrier-mediated transport of rC **3** and araC **5** on the concentration of sodium tetraborate **9** in the aqueous source phase. In these experiments, the source phase contained a 1:1 mixture of rC **3** and araC **5** (25 mM) and, as indicated, the concentration of sodium tetraborate **9** added to the source phase was varied between 0-25 mM. The organic CHCl_3 phase contained G **1** (20 mM). Aliquots from the receiving phase were assessed at multiple time points; the data in the above graph indicates the relative ratios of transport of rC **3** and araC **5** after 16 hours. The higher the concentration of sodium tetraborate **9** in the source phase, the higher the ratio of araC **5** to rC **3** in the receiving phase.

Figure S12. Reinitiation of carrier-mediated transport of rC **3** transport in the presence of $\text{Na}_2\text{B}_4\text{O}_7$ upon addition of a competing diol. In these experiments, the source phase initially contained a 1:1 mixture of rC **3** and araC **5** (25 mM) and sodium tetraborate **9** (25 mM). The organic CHCl_3 phase contained G **1** (20 mM). The receiving phase was monitored by NMR at different time periods. As seen previously, from time 0-24 h, a preference is seen for transport of araC **5** (squares) over rC **3** (diamonds). At $t=24$ h, an excess of ribose **10** (0.5 M) was added to the source phase. The NMR of the receiving phase shows that transport of rC **3** transport was reinitiated after addition of ribose **10**. Ribose **10** competes with rC **3** for chelation of the boron, resulting in dissociation of the cytidine borate esters **11-13** in the source phase, and release of “free” rC **3** for subsequent transport.

Figure S13. DOSY spectra for borate esters **11-13**. Diffusion experiments were carried out on a Bruker AVIII-600, using the Stimulated Echo Pulse Gradient sequence in FT mode. All samples for the diffusion measurements were prepared in Shigemi tubes (Shigemi, Inc., Allison Park, PA) and the temperature was actively controlled at 25.0 °C. Diffusion coefficients were derived using integration of the desired peaks to a single exponential decay, using the “Simfit (Bruker XWINNMR v3.1)” software. Experiments consisted of 32 points and comprised 100 scans with a pulse delay of 4 s, δ value of 2.1 ms, and Δ value of 60.0 ms. The samples were prepared in a 5 to 1 ratio of rC **3** (75mM) to sodium tetraborate (12.5 mM) in D₂O. The diffusion coefficients for **11-13** are shown in Table S2.

Table 1. ^1H NMR chemical shift assignments (as determined by 2D COSY) for cytidine borate esters **11-13** recorded on a Bruker AVIII-600 operating at 600.13 MHz.

	Monomer 11 δ (ppm)	Dimer (a) 12/13 δ (ppm)	Dimer (b) 12/13 δ (ppm)
H6	7.67-7.70*		
H5	6.04-6.06*		
H1'	5.88 (d)	5.83 (d)	5.82 (d)
H2'	4.53 (dd)	4.62 (dd)	4.56 (dd)
H3'	4.40 (dd)	4.45 (dd)	4.47 (dd)
H4'	4.14 (m)	4.08 (m)	4.12 (m)
H5',5''	3.77-3.96 (m)*		

* These peaks were unable to be accurately assigned due to overlap.

Table 2. Diffusion coefficients calculated for cytidine borate esters **11-13** at 25.0 °C.

	D ($10^{-10} \text{ m}^2 \text{ s}^{-1}$)
rC 3	4.708 ± 0.056
Monomer 11	4.244 ± 0.102
Dimer 12/13	3.376 ± 0.061

References

[1] (a) Marlow, A. L.; Mezzina, E.; Spada, G. P.; Masiero, S.; Davis, J. T.; Gottarelli, G. *J. Org. Chem.* **1999**, *64*, 5116-5123. (b) Shi, X.; Mullaugh, K. M.; Fettingner, J. C.; Jiang, Y.; Hofstadler, S. A.; Davis, J. T. *J. Am. Chem. Soc.* **2003**, *125*, 10830-10841.

[2] Williams, N. G.; Williams, L. D.; Shaw, B. R. *J. Am. Chem. Soc.* **1989**, *111*, 7205-7209.