Enhancing Hoogsteen Interactions: A Pyrrole-Containing Purine Nucleoside that Competes with Guanosine Self-Assembly

Jonathan L. Sessler*, Janarthanan Jayawickramarajah, Courtney L. Sherman, and Jennifer S. Brodbelt

Department of Chemistry and Biochemistry and Institute for Cellular and Molecular Biology, 1 University Station A-5300, The University of Texas at Austin, Austin, Texas, 78712-0165

Supporting Information

I. Synthetic experimental

II. Variable temperature $^1$H-NMR studies

III. $^1$H-NMR titrations:
   IIIa. Titrations using guanosine 10 as host
   IIIb. Titrations using nucleobase 1 as host
   IIIc. Dimer-monomer self-association of 10

IV. Electrospray mass-spectrometric studies

V. ITC titration studies
I. Synthetic experimental

Nucleobases 3, 7, and 10 were synthesized using a standard silylating protocol.\textsuperscript{1,2} Guanosine 11 was purchased from Sigma. \textit{tert}-Butoxycarbonyl protected pyrrole boronic acid 9 was synthesized using a literature procedure.\textsuperscript{3} The protocol for alkylation of the imino NH to the N-propyl derivative was derived from a previous literature procedure.\textsuperscript{4} The Suzuki cross-coupling reaction of 9 and the subsequent deprotection method were worked out by modifying literature protocols.\textsuperscript{3,5}
1-Propyl-9-[2′,3′,5′-tri-O-(tert-butyldimethylsilyl)-β-D-ribofuranosidyl]-purin-6-one (8): To a solution of dry DMF (35 ml), under an argon atmosphere was added 7 (2.0 g, 3.28 mmol) and DBU (0.51 ml, 3.44 mmol). To the resulting solution was added 1-bromopropane (0.30 ml, 3.28 mmol) dropwise. The reaction was stirred at room temperature for 24 hr. After evaporation of solvent, the resulting crude product was purified by column chromatography on a silica-gel column utilizing ethyl acetate/hexanes (1/1) as the eluent. This yielded 8 as a white powder (1.17 g, 55%). \(^1^H\) NMR (500 MHz, CDCl\(_3\)) \(\delta\) 8.09 (s, 1 H), 7.89 (s, 1 H), 5.92 (d, \(J = 5.5\) Hz, 1 H), 4.44 (dd, \(J = 5.5, 1.0\) Hz, 1 H), 4.24 (t, \(J = 4\) Hz, 1 H), 4.09-3.97 (m, 3 H), 3.92 (dd, \(J = 11.4, 3.0\) Hz, 1 H), 3.74 (dd, \(J = 11.4, 3.0\) Hz, 1 H), 1.80-1.75 (m, 2 H), 0.93-0.87 (m, 21 H), 0.75 (s, 9 H), 0.09 (s, 3 H), 0.08 (s, 3 H), 0.054 (s, 3 H), 0.047 (s, 3 H), -0.08 (s, 3 H), -0.25 (s, 3 H); \(^1^3^C\) NMR (125 MHz, CDCl\(_3\)) \(\delta\) 156.5, 147.4, 146.8, 138.5, 124.8, 87.9, 85.5, 76.5, 71.9, 62.5, 48.3, 26.0, 25.8, 25.6, 22.9, 18.4, 18.0, 17.8, 10.9, -4.5, -4.7, -4.8, -5.2, -5.46, -5.48; HR-MS (CI): \(m/z\) 653.3930 [M\(^+\)], calcd for C\(_{31}\)H\(_{60}\)N\(_4\)O\(_5\)Si\(_3\): 653.3950.
8-Bromo-1-propyl-9-[2',3',5'-tri-O-(tert-butyldimethylsilyl)-ß-D-ribofuranosidyl]-purin-6-one (4): To a solution of dry DMF (25 ml) under an argon atmosphere was added 3 (0.87 g, 1.26 mmol) and DBU (0.20 ml, 1.32 mmol). To this solution was added dropwise 1-bromopropane (0.11 ml, 1.26 mmol). After 17 hr, TLC analysis indicated the presence of a less polar spot (30% ethyl acetate in hexanes, eluent). The DMF was then removed in vacuo, and DCM was added to the resulting residue. The organic layer was washed twice with sodium bicarbonate, dried over Na$_2$SO$_4$, and the resulting crude product was chromatographed on a silica-gel column (30% ethyl acetate in hexanes). The isolated yield of 4 was (0.70 g, 76%). $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ 7.84 (s, 1 H), 5.89 (d, $J$ = 6.0 Hz, 1 H), 5.17 (dd, $J$ = 6.0, 4.5 Hz, 1 H), 4.43 (dd, $J$ = 4.5, 3.0 Hz, 1 H), 4.01-3.87 (m, 4 H), 3.67 (dd, $J$ = 11.0, 4.5 Hz, 1 H), 1.80-1.70 (m, 2 H), 0.91-0.89 (m, 12 H), 0.78 (s, 9 H), 0.73 (s, 9 H), 0.08 (s, 3 H), 0.07 (s, 3 H), -0.04 (s, 3 H), -0.08 (s, 3 H), -0.12 (s, 3 H), -0.36 (s, 3 H); $^{13}$C NMR (125 MHz, CDCl$_3$) $\delta$ 155.1, 148.3, 146.4, 127.0, 125.6, 90.3, 85.5, 72.2, 71.9, 62.1, 48.5, 25.8, 25.7, 25.5, 22.8, 18.2, 17.9, 17.7, 10.8, -4.6, -4.67, -4.70, -5.3, -5.5, -5.6; HR-MS (CI): m/z 731.3068 [M$^+$], calcd for C$_{31}$H$_{59}$N$_4$O$_5$Si$_3$Br: 731.3055.
8-(1″-tert-Butoxycarbonyl-pyrrol-2″yl)-1-propyl-9-[2′,3′,5′-tri-O-(tert-
butyldimethylsilyl)-β-D-ribofuranosidyl]-purin-6-one (5): To a mixture of 4 (0.70 g, 0.96 mmol), and tert-butoxycarbonyl (BOC) protected pyrroleboronic acid 9 (0.34 g, 1.44 mmol) was added DME (15 ml). The resulting solution was bubbled with argon for 30 min, after which tetrakis(tri-phenylphosphine) palladium (0.06 g, 5 mol %) was added. A solution of Na₂CO₃ (0.53 g, 5.0 mmol) in 2 ml of de-gassed H₂O was added to the reaction mixture. Dry toluene (6 ml) was also added as a co-solvent. The reaction was stirred and heated at 85 ºC for 24 hr. After cooling, the solvents were removed in vacuo and DCM was added to the resulting residue. The organic layer was washed twice with a saturated solution of sodium bicarbonate, and dried over Na₂SO₄. After evaporation of the solvent, the crude compound was purified via silica-gel column chromatography (30% ethyl acetate in hexanes) yielding pure 5 (0.45 g, 58%). ¹H NMR (500 MHz, CDCl₃) δ 7.86 (s, 1 H), 7.40 (dd, J = 3.5, 1.8 Hz, 1 H), 6.53 (dd, J = 3.5, 1.5 Hz, 1 H), 6.24 (t, J = 3.5 Hz, 1 H), 5.67 (d, J = 6.0 Hz, 1 H), 4.91(dd, J = 6.0, 4.5 Hz, 1 H), 4.26 (dd, J = 4.5, 2.3 Hz, 1 H), 4.04-3.96 (m, 2 H), 3.90-3.87 (m, 1 H), 3.62-3.53 (m, 2H), 1.82-1.78 (m, 2 H), 0.94 (t, J = 7.5 Hz, 3 H), 0.85 (s, 9 H), 0.84 (s, 9 H), 0.70 (s, 9 H), 0.06 (s, 3 H), 0.04
(s, 3 H), 0.01 (s, 3 H), -0.01 (s, 3 H), -0.14 (s, 3 H), -0.36 (s, 3 H); $^{13}$C NMR (125 MHz, CDCl$_3$) δ 156.3, 148.1, 147.8, 146.1, 144.7, 124.5, 123.6, 120.7, 119.1, 110.6, 88.3, 84.9, 84.3, 72.9, 72.3, 62.7, 48.3, 27.4, 25.9, 25.8, 25.6, 23.0, 18.3, 18.0, 17.8, 10.8, -4.4, -4.6, -4.7, -5.2, -5.4, -5.5; HR-MS (CI): $m/z$ 818.4741 [M$^+$], calcd for C$_{40}$H$_{71}$N$_5$O$_7$Si$_3$: 818.4740; calcd for C$_{40}$H$_{71}$N$_5$O$_7$Si$_3$: C 58.71, H 8.75, N 8.56; found C 58.79, H 8.90, N 8.47.

1-Propyl-8-(2′′-pyrrolyl)-9-[2′,3′,5′-tri-O-(tert-butyldimethylsilyl)-β-D-ribofuranosidyl]-purin-6-one (1): To a round bottom flask containing 5 (0.09 g, 0.11 mmol) under an argon atmosphere was added dry THF (2 ml). A solution of sodium methoxide (0.02 g, 0.44 mmol) in 1 ml of dry methanol was added to the reaction mixture. The reaction was left to stir at room temperature for 3 hr, upon which time the reaction was quenched with H$_2$O. To the resulting solution was added ethyl acetate and the organic layer was washed twice with H$_2$O and once with brine. After drying over Na$_2$SO$_4$, the resulting residue was purified on a silica-gel column (30% ethyl acetate in hexanes) to give 1 (0.05 g, 63%). $^1$H NMR (500 MHz, CDCl$_3$) δ 9.56 (bs, 1H), 7.81 (s, 1 H), 6.95-6.94 (m, 1 H), 6.83-6.81 (m, 1 H), 6.30-6.27 (m, 2 H), 5.31 (dd, $J = 6.0, 4.5$ Hz, 1 H), 4.55 (dd, $J = 4.5, 3.0$ Hz, 1 H), 4.09-3.93 (m, 4 H), 3.73 (dd, $J = 11.0, 4.5$ Hz, 1 H), 1.89-1.76 (m, 2 H),
0.97 (t, J = 7.5 H, 3 H), 0.95 (s, 9 H), 0.82 (s, 9 H), 0.73 (s, 9 H), 0.14 (s, 6 H), 0.00 (s, 3 H), -0.06 (s, 3 H), -0.09 (s, 3 H), -0.31 (s, 3 H); $^{13}$C NMR (125 MHz, CDCl$_3$) $\delta$ 156.5, 148.0, 146.0, 145.1, 125.0, 121.2, 120.6, 111.5, 110.2, 88.7, 85.1, 72.8, 72.1, 62.3, 48.4, 25.86, 25.80, 25.7, 23.1, 18.2, 18.1, 17.9, 11.0, -4.47, -4.52, -4.6, -5.1, -5.4, -5.5; HR-MS (Cl): $m/z$ 718.4228 [M$^+$], calcd for C$_{35}$H$_{63}$N$_5$O$_5$Si$_3$: 718.4215; calcd for C$_{35}$H$_{63}$N$_5$O$_5$Si$_3$: C 58.53, H 8.84, N 9.75; found C 58.61, H 8.93, N 9.71.
II. Variable temperature $^1$H-NMR studies

Variable temperature experiments ($^1$H-NMR in CD$_2$Cl$_2$) were carried out using a INOVA Varian 500 MHz NMR spectrometer. The splitting of the guanosine 10 imino NH resonance into two signals is clear at low temperatures (Figure 1). The two resonances correspond to the GG dimer (12.7 ppm) and the proposed ensemble I (12.2 ppm).

![Figure 1. $^1$H-NMR spectra of the downfield region of a mixture of 1 and 10 in CD$_2$Cl$_2$,](image-url)
recorded at varying temperatures: (a) 27 °C, (b) -20 °C, (c) -40 °C, (d) -60 °C, and (e) -80 °C. [1] = 56 mM, [10] = 32 mM.

\[ \text{Evidence for dimerization of 10 was obtained from} \] 

\[ \text{\textsuperscript{1}H-NMR spectroscopy. In} \]

\[ \text{particular, the imino NH proton present in 10 at 27 °C is seen to resonate at 12.3 ppm} \]

\[ \text{(Figure 2), Such a downfield shift is consistent with a strongly hydrogen bonded species,} \]

\[ \text{proposed to be the GG dimer.} \]
Figure 2. Portion of the $^1$H-NMR spectrum of 10 in CD$_2$Cl$_2$ at 27 °C. [10] = 32 mM

When the $^1$H-NMR spectrum of 10 is recorded in CD$_2$Cl$_2$ at -80 °C (Figure 3) one imino NH proton signal at 12.7 ppm is seen. Such an observation is consistent with the GG dimer being the dominant species under these conditions.
Figure 3. Portion of the $^1$H-NMR spectrum of 10 in CD$_2$Cl$_2$ recorded at -80 °C. [10] = 32 mM.

III. $^1$H-NMR titrations

All $^1$H-NMR-based titration studies were carried out at room temperature using a Varian Unity Plus 400 MHz NMR spectrometer. Host and guest samples were dried
overnight in a vacuum oven at 35 ºC. CDCl₃ was purchased from Cambridge Isotope Laboratories, Inc, and passed through a plug of basic alumina (50-200 µm) before use.

IIIa. ¹H-NMR titrations using guanosine 10 as host

The host (10) concentration was kept constant through each titration. In order to circumvent dilution effects, the guest solutions used to effect the titration contained equimolar concentration of host as the host solution.

¹H-NMR titrations were carried out with 10 as the host using the change in the imino proton chemical shift to monitor the presumed binding event. However, when 1, 5, and 8 were added to 10 no downfield shifts were evident. Rather, small upfield shifts were observed. These shifts (Figure 4a) can be explained by the aggregation of 10 to form a GG dimer (see reference 10 in main text) in the absence of an added guest. Upon addition of a guest, such as 1, 5, or 8, there is a break up of this dimer, resulting in a slight upfield shift in the imino proton signal. Consistent with such an explanation, the addition of nucleoside 1 induced the largest upfield shift, leading us to infer that this compound competes more effectively with the GG dimerization process than either 5 or 8.
Figure 4a. $^1$H-NMR titrations for 10 upon addition of 1 (●), 5 (×), and 8 (□), when the imino NH resonance of 10 is monitored.

When the amino NH$_2$ resonance of 10 is followed a significant downfield shift was evident upon addition of increasing concentrations of nucleobase 1. In contrast, controls 8 and 5 only show a small downfield shift (Figure 4b). These findings support the hypothesis that nucleobase 1 with the proposed three-point Hoogsteen interaction can bind guanosine 10 with higher affinity than control compounds 5 and 8.
Figure 4b. $^1$H-NMR binding titrations for 10 upon addition of 1 (●), 5 (▲), and 8 (□), when the amino NH$_2$ resonance of 10 is used to monitor binding.

IIIb. $^1$H-NMR titrations using 1 as Host

$^1$H-NMR titrations were carried out using 1 as the host. The pyrrolic NH proton signal was followed as the concentration of 10 was increased (Figure 5). The host (1) concentration was kept constant at 1.4 mM. In order to calculate the binding constant ($K_a$)
between 1 and 10 the free concentration of 10 ([G]) after each injection was first determined. This was achieved by substituting eqn. 1 (where \( K_{\text{dimer}} = 4.7 \pm 0.8 \times 10^2 \text{ M}^{-1} \)), determined experimentally (see part IIIc)) into eqn. 2 (where \( C_G \) is the formal concentration of 10, [G] is the free concentration of 10, and \([GG]\) is the concentration of dimer) resulting in quadratic eqn. 3. Eqn. 3 is solved using known values for \( C_G \) to determine values for [G].

\[
K_{\text{dimer}} = \frac{[GG]}{[G]^2} \quad (1)
\]

\[
C_G = [G] + 2[GG] \quad (2)
\]

\[
2K_{\text{dimer}}[G]^2 + [G] - C_G = 0 \quad (3)
\]
**Figure 5.** $^1$H-NMR titration curve derived from the titration of compound 1 and guanosine 10 in CDCl$_3$. Note the x-axis depicts the concentration of free 10. The curved line shows the fit to a 1:1 binding profile as described by Connors.$^6$

**IIIc. Monomer-Dimer self-association of 10**

The monomer-dimer self-association of 10 using $^1$H-NMR dilution studies in CDCl$_3$ has been previously studied by Schneider and coworkers, by following the amino NH$_2$ resonance.$^7$ In our hands, $^1$H-NMR dilution studies in the same solvent resulted in a $K_{dimer}$ of $4.68 \pm 0.8 \times 10^2$ M$^{-1}$ (a value that is close to that determined by Schneider of $3 \times 10^2$ M$^{-1}$).
Figure 6. $^1$H-NMR self-association curve of guanosine 10 in CDCl$_3$: [10] = 5.9 X $10^{-4}$ M to 3.3 X $10^{-2}$ M. The curved line is the fit to a monomer-dimer self-association profile.\(^8\)

IV. Electrospray mass-spectrometric studies

Electrospray ionization mass spectrometric analysis of solutions containing 1, 5, 8 and 11 in methanol at concentrations of 10 µM were performed on a ThermoFinnigan (San Jose, CA) LCQ duo quadrupole ion trap mass spectrometer. The instrument was tuned to minimize in-source fragmentation, and the heated capillary was held at 80 °C. Solutions were introduced to the electrospray ion source via a syringe pump (Harvard Apparatus PHD 2000, South Natick, MA) with a flow rate of 5 µL/min. The spectrum of 11 was analyzed alone and in the form of 1:1 mixtures of the other three compound
Figure 7. ESI mass spectra of 11 alone (a) and with 1 (b), 5 (c), and 8 (d). On addition of 1 to 11 (b), the abundances of [(11)₂ + Na]⁺ and [(11)₈ + 2•Na]²⁺ decreased significantly and the peak ascribed to [1 + 11 + Na]⁺ was observed. Upon the addition of 5 to 11 (c), neither change was observed. On addition of 8 to 11 (d), the abundance of [(11)₈ + 2•Na]²⁺ decreased, and [8 + 11 + Na]⁺ was observed, but to a lesser extent than what is seen for mixtures of 1 and 11. Presumably, this reflected a reduced interaction between 8 and 11.
V. ITC titration studies

Va. Titrations of 1, 5, 8, with 10

Microcalorimetric titrations were performed using an Isothermal Titration Calorimeter purchased from Microcal Inc., MA. The experimental temperature was kept at 30 °C. The ORIGIN software provided by Microcal Inc. was used to calculate binding constant \( K_a \). Spectral grade 1,2 dichloroethane (ACROS) was used as purchased. In all cases, guanosine 10 was used as the guest and 1, 5, and 8 were used as the hosts (See Figure 8a. All the ITC sample titration data were subtracted from reference titrations involving the addition of 10 into pure solvent (1,2 dichloroethane). This subtraction is expected to eliminate any effect due to the heat of dilution, heat of dissociation of GG dimeric or oligomeric species, and account for any systematic errors involving the injection of guest into the solvent.

The ITC results provide support for the notion that control compound 5 does not bind 10 appreciably in solution under these conditions. Control system 8 displays evidence of some slight binding. However, the interaction is too weak to allow for an accurate fit. This differs dramatically to what is seen in the case of 1 and 10. Here, the resulting ensemble (I) is held together quite well, being characterized by an estimated binding constant \( K_a \) of 8.5 ± 0.6 x 10^3 M^{-1} (Figure 8b).
**Figure 8a.** ITC binding curves (normalized integration data after subtraction from the reference titration given in terms of Cal/mole of guest (10) injected) for 1 (●), 5 (×), and 8 (□) as generated from experiments involving addition of 10. For titration of 10 with 1: [10] = 0-1.12 mM and [1] = 0.63-0.56 mM. For titration of 10 with 5: [10] = 0-1.11 mM and [5] = 0.63-0.55 mM. For titration of 10 with 8: [10] = 0-1.10 mM and [8] = 0.56-0.50 mM.
Figure 8b. Compound 1 titrated with 10 as studied by ITC. [10] = 0-1.12 mM and [1] = 0.63-0.56 mM. The raw ITC data involves µcal/second plotted against time in minutes. After the integration baseline has been subtracted (top), the normalized integration data after subtraction from the reference titration (bottom) in terms of kcal/mole of injectant (10), is shown. The solid line shows the computer generated curve fit from which the binding constant is inferred.

Vb. Dimer-monomer self-association of 10

Guanosine 10 dimer dissociation

The injection of guanosine 10 into pure solvent (1,2 dichloroethane) resulted in a characteristic dimer-dissociation curve (See Figure 9, dotted line). This titration was then applied to a model of homomeric binding as reported by Krische and coworkers. This computation resulted in a dissociation constant of $3.5 \pm 0.2 \times 10^{-4}$ M, from which an association constant of $2.9 \pm 0.2 \times 10^3$ M$^{-1}$ was calculated.
**Figure 9.** ITC derived dimer dissociation of 10 in 1,2 dichloroethane. [10] in the cell = 0-1.12 mM. The dotted line is the experimental data and the solid line is the generated curve fit.
References


