Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2020

Supporting Information for:

Design, Synthesis and Structure of a Frustrated Benzoxaborole and its Applications in the Complexation of Amines, Amino Acids, and Protein Modification

Jasmine Bhangu, Randy M. Whittal, Dennis G. Hall*

Department of Chemistry, 4-010 Centennial Center for Interdisciplinary Science, University of Alberta,

Edmonton, Alberta T6G 2G2, Canada

dennis.hall@ualberta.ca

Table of Contents

1.	General Information			
2.	Chemical Synthesis and Analytical Data			
3.	Molecular Modeling (Figure 2b)			
4.	High Resolution ESI-MS Spectra			
4	.1 Compound 2			
4.2 Benzoxaborole				
4.3 Compound 12				
5.	NMR pKa Titration Data			
5	.1 Benzoxaborole Acid/Base Titration Curve			
6.				
6	5.1 Compound 2 Lysozyme Conjugation MS (Figure 12)			
6	5.2 2-FBPA Lysozyme Conjugation MS			
6	5.3 2-APBA Lysozyme Conjugation MS			
6.4 Compound 2 and L-Cysteine Conjugate: NMR and MS Spectra				
6.5 LC-MS of Compound 2 Conjugation with Benzylamine				
6	6.6 LC-MS of Compound 2 Conjugation with 2-Hydroxybenzylamine			
7.	References			
8.	NMR Spectra			

1. General Information

Unless otherwise indicated, all reactions were performed under a nitrogen atmosphere using glassware that was washed thoroughly prior to use. All reagents were purchased from Sigma-Aldrich, Combi-Blocks or Alfa Aesar and used as received except diisopropylamine and benzylamine, which were distilled over calcium hydride prior to use. DMF and THF were used directly from an MBraun Solvent Purification System. Thin layer chromatography (TLC) was performed on Merck Silica Gel 60 F254 plates and visualized with UV light and *p*-anisaldehyde stain. Flash chromatography was performed on ultra-pure silica gel 230-400 mesh. Nuclear magnetic resonance (NMR) spectra were recorded on Agilent/Varian INOVA-400 and INOVA-500 MHz instruments. The residual solvent protons (¹H) of CDCl₃ (7.26 ppm), Acetone-d₆ (2.05 ppm), DMSO-d₆ (2.50 ppm), and D₂O (4.79 ppm) were used as internal standards, and the carbons signal (¹³C) of CDCl₃ (77.06 ppm), DMSO-d₆ (39.51) and acetone-d₆ (29.84 and 206.26 ppm) were used as an internal standard. MestReNova software was used to analyze all of the NMR data. The following abbreviations are used in reporting NMR data: s, singlet; d, doublet; t, triplet; app t, apparent triplet; q, quartet; dd, doublet of doublets; ddd, doublet of doublet of doublets; ddtd, triplet of doublet; m, multiplet; comp m, complex multiplet. In ¹³C NMR spectroscopy, the quaternary carbon bound to the boron atom is often missing due to the quadrupolar relaxation of boron. This effect was observed in each boronic acid or boronate ester compound. Infrared spectra (performed on a Nicolet Magna-IR 750 instrument equipped with a Nic-Plan microscope) were recorded by the University of Alberta Analytical and Instrumentation Laboratory. High-resolution mass spectra were recorded by the University of Alberta mass spectrometry services laboratory using either electron impact (EI) or electrospray ionization (ESI) techniques. High-resolution mass spectra (HRMS) using electrospray ionization (ESI), and

LC-MS mass spectra were recorded by the University of Alberta Mass Spectrometry Services Laboratory. ESI-FTICR-MS was performed on Bruker Daltonics 9.4T Apex-Qe FTICR MS with Apollo II Dual source. Positive ion electrospray was used and the instrument was tuned to maximize sensitivity over the m/z range 900 to 2500. The instrument was calibrated with Agilent low concentration tuning mix.

2. Chemical Synthesis and Analytical Data



2,6-dibromobenzaldehyde (6): To a flame-dried 250 mL round bottom flask under nitrogen was added dry THF (50.0 mL) and diisopropylamine (1.30 mL, 9.28 mmol) and cooled to – 78 °C. To the solution was added *n*-BuLi (3.80 mL, 9.50 mmol) dropwise and the reaction was stirred for 15 min. 1,3-dibromobenzene **5** (0.96 mL, 7.94 mmol) was added dropwise at a rate of ~1 drop per second and the solution was allowed to stir for 30 min before DMF (1.20 mL, 15.5 mmol) was added. The reaction was allowed to warm to room temperature and stopped by addition of 2 M H₂SO₄ (10 mL), diluted with EtOAc (20 mL) and the layers separated. The aqueous layer was extracted with EtOAc (2 × 10 mL) and the combined organic extracts were washed with brine, dried over Na₂SO₄, filtered and concentrated *in vacuo* to afford the title compound as a pale-yellow solid (1.90 g, 91%) without any further purification. All spectral data matched the literature.¹

¹H NMR (498 MHz, CDCl₃) δ 10.26 (s, 1H), 7.64 (d, J = 8.0 Hz, 2H), 7.22 (t, J = 8.0 Hz, 1H).
¹³C NMR (126 MHz, CDCl₃) δ 191.3, 134.1, 133.8, 133.1, 125.0.

IR (cast film, cm⁻¹): 3386, 3360, 3139, 3084, 2889, 2774, 1701, 1661, 1571, 1551, 1429, 1400, 1270, 1204, 1184, 1070, 1059.

HRMS (EI) for C₇H₄O⁸¹Br₂: *calcd*.: 265.8588; *found*: 265.8585.



1,3-dibromo-2-(dimethoxymethyl)benzene (7): To a 250 mL round bottom flask was added **6** (1.90 g, 7.20 mmol), pTSA•H₂O (164 mg, 0.864 mmol) and trimethyl orthoformate (4.05 mL, 36.9 mmol) and MeOH (60 mL). The reaction was stirred at room temperature for 18 h under nitrogen then stopped by addition of NaHCO₃ and diluted with water. The layers were separated and the aqueous layer was extracted with EtOAc (2 × 10 mL) and the combined organic extracts washed with brine, dried over Na₂SO₄, filtered and concentrated *in vacuo* to afford the title compound as a yellow solid (1.93 g, 86%) without any further purification. All spectral data matched the literature.¹

¹**H NMR** (498 MHz, CDCl₃) δ 7.56 (d, *J* = 8.0 Hz, 2H), 7.01 (t, *J* = 8.0 Hz, 1H), 5.83 (s, 1H), 3.48 (s, 6H).

¹³C NMR (126 MHz, CDCl₃) δ 135.2, 133.7, 130.8, 124.1, 107.0, 56.0.

IR (cast film, cm⁻¹): 3073, 3001, 2933, 2903, 2834, 1570, 1553, 1444, 1380, 1363, 1215, 1207, 1188, 1148, 1105, 1075.

HRMS (EI) for C₉H₁₀O₂⁷⁹Br⁸¹Br: *calcd*.: 309.9027; *found*: 309.9028.



3-bromo-2-(dimethoxymethyl)benzaldehyde (8): To a flame-dried 50 mL round bottom flask under nitrogen was added **7** (1.85 g, 5.97 mmol) and dissolved in dry THF (10 mL) and cooled to - 78 °C before *n*-BuLi (2.39 mL, 5.97 mmol) was added drop-wise and the reaction was stirred for 15 min. To the reaction was added DMF (600 µL, 7.76 mmol) and stirred for 30 min. The reaction was allowed to warm to room temperature and stopped with addition of 1 M HCl (3 mL), diluted with EtOAc (20 mL) and the layers separated. The aqueous layer was extracted with EtOAc (3 × 10 mL) and the combined organic extracts washed with brine, dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography (5:1 Hexanes:EtOAc) to afford the title compound as a yellow crystalline solid (1.42 g, 92%). All spectral data matched the literature.¹

¹H NMR (498 MHz, CDCl₃) δ 10.69 (d, J = 0.9 Hz, 1H), 7.89 (dd, J = 7.7, 1.3 Hz, 1H), 7.73 (dd, J = 8.0, 1.3 Hz, 1H), 7.29 (td, J = 7.8, 0.9 Hz, 1H), 5.86 (s, 1H), 3.51 (s, 6H).
¹³C NMR (126 MHz, CDCl₃) δ 192.6, 139.4, 137.8, 137.2, 130.4, 128.0, 123.5, 107.1, 56.3.

IR (cast film, cm⁻¹): 3072, 2998, 2934, 2912, 2832, 1692, 1586, 1566, 1447, 1376, 1241, 1211, 1121, 1104, 1065.

HRMS (EI) for C₁₀H₁₁O₃⁸¹Br: *calcd*.: 259.9871; *found*: 259.9859.



2-(dimethoxymethyl)-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzaldehyde (9): To a flame-dried 250 mL round bottom flask under argon was added **8** (1.58 g, 6.11 mmol), B₂pin₂ (3.10 g, 12.2 mmol), KOAc (1.80 g, 18.3 mmol) and Pd(dppf)Cl₂•DCM (500 mg, 0.611 mmol) and freshly distilled dioxane (60 mL). The solution was degassed with argon for 10 mins then heated to 80 °C and stirred for 18 h. The reaction was brought to room temperature and filtered through Celite and the filtrate collected and washed with a saturated aqueous solution of ammonium chloride. The organic extracts were concentrated in vacuo and purified by flash chromatography (40:1 DCM:EtOAc) to afford the title compound as a yellow solid (1.00 g, 53%).

¹**H** NMR (498 MHz, CDCl₃) δ 10.64 (s, 1H), 7.98 (dd, J = 7.7, 1.5 Hz, 1H), 7.85 (dd, J = 7.4, 1.5 Hz, 1H), 7.44 (td, J = 7.6, 0.8 Hz, 1H), 6.12 (s, 1H), 3.43 (s, 6H), 1.38 (s, 12H). Excess B₂pin₂ seen at around 1.25 ppm in spectrum.

¹³C NMR (126 MHz, CDCl₃) δ 193.9, 145.5, 139.2, 135.0, 131.1, 128.3, 104.4, 84.3, 83.7, 55.3, 25.2, 25.1.

¹¹**B** NMR (160 MHz, CDCl₃) δ 30.7.

IR (cast film, cm⁻¹): 2979, 2932, 2832, 1688, 1578, 1473, 1380, 1373, 1350, 1127, 1069. **HRMS (EI)** for $C_{16}H_{22}O_5^{11}B(M + H)^+$: *calcd*.: 305.1560; *found*: 305.1561.



(2-(dimethoxymethyl)-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)methanol (10):

To a 50 mL round bottom flask under nitrogen was added 9 (1.00 g, 3.86 mmol) and dissolved in

MeOH (12 mL) and cooled to 0 °C before NaBH₄ (219 mg, 5.79 mmol) was added portion-wise and the reaction stirred for 1 h. The reaction was stopped with addition of 5 mL water and about half of the MeOH was removed *in vacuo*. The solution was diluted with EtOAc, the phases separated, and the aqueous phase extracted with EtOAc (3 × 15 mL). The combined organic extracts were washed with brine, dried over Na₂SO₄, filtered and concentrated *in vacuo*. The crude material was purified by flash chromatography (3:1 Hexanes:EtOAc) to afford the title compound as colourless oil (857 mg, 84%).

¹**H NMR** (498 MHz, CDCl₃) δ 7.59 (dd, *J* = 7.4, 1.5 Hz, 1H), 7.42 (dd, *J* = 7.5, 1.5 Hz, 1H), 7.32 (t, *J* = 7.5 Hz, 1H), 5.96 (s, 1H), 4.75 (d, *J* = 6.7 Hz, 2H), 3.43 (s, 6H), 1.37 (s, 12H).

¹³C NMR (126 MHz, CDCl₃) δ 140.9, 139.4, 133. 8, 133.0, 128.5, 105.1, 83.8, 64.2, 55.0, 25.1.
¹¹B NMR (160 MHz, CDCl₃) δ 30.7.

IR (cast film, cm⁻¹): 3450, 3063, 2933, 2832,1591, 1474, 1444, 1371, 1346, 1292, 1170, 1148, 1131.

HRMS (ESI-TOF): for $C_{16}H_{25}BO_5Na (M + Na)^+$: *calcd*.: 331.1687; *found*: 331.1690.



potassium trifluoro(3-hydroxy-1,3-dihydroisobenzofuran-4-yl)borate (11): To a 25 mL round bottom flask was added 10 (423 mg, 1.37 mmol) and dissolved in MeOH (4.15 mL) followed by the addition of 4.5 M aqueous KHF_2 (1.67 mL, 7.54 mmol). The reaction was stirred for 10 min then immediately concentrated *in vacuo*. The solid white material was filtered through Celite using a solution of hot 8% MeOH in acetone and the filtrate collected and

concentrated *in vacuo*. The crude material was dissolved in a minimal amount of hot acetone and recrystalized with diethyl ether to afford the title compound as a white solid (209 mg, 63%) without any further purification.

compound as colourless oil (857 mg, 84%).

¹**H NMR** (498 MHz, D₂O) δ 7.56 (d, *J* = 7.3 Hz, 1H), 7.44 (t, *J* = 7.4 Hz, 1H), 7.34 (d, *J* = 7.5 Hz, 1H), 6.65 (s, 1H), 5.27 (d, *J* = 12.5 Hz, 1H), 5.02 (d, *J* = 12.7 Hz, 1H).

¹³C NMR (126 MHz, D₂O) δ 141.3, 138.6, 132.2, 132.2, 129.6, 121.1, 102.1, 71.8.

¹¹**B** NMR (160 MHz, D₂O) δ 3.32.

IR (cast film, cm⁻¹): 3524, 3053, 3017, 2947, 2903, 2859, 1434, 1413, 1244, 1231, 1069, 1044, 1010, 977.

HRMS (ESI-TOF): for C₈H₇F₃BO₅ (M)⁻: *calcd*.: 203.0491; *found*: 203.0492.



(3-hydroxy-1,3-dihydroisobenzofuran-4-yl)boronic acid (2-I): To a 10 mL round bottom flask was added 2-17 (80 mg, 0.331 mmol) and silica gel (20.0 mg, 0.331 mmol) and water (1.10 mL) and the reaction was stirred for 4 h at rt. The reaction was diluted with acetone (8 mL) and the silica gel filtered off and the filtrate concentrated *in vacuo*. The solid white material was dissolved in water and any intractable organic solids were filtered. The filtrate was collected and concentrated *in vacuo* to afford the title compound as a white solid (60.3 mg, 85%) without any further purification.

¹**H** NMR (498 MHz, Acetone-d₆ + 10 μL H₂O) δ 7.77 (dd, *J* = 5.7, 2.8 Hz, 1H), 7.63 (s, 2H), 7.38 (d, *J* = 5.7 Hz, 2H), 6.70 (d, *J* = 7.1 Hz, 1H), 6.56 (dd, *J* = 4.5, 2.4 Hz, 1H), 5.18 (dd, *J* = 12.7, 2.4 Hz, 1H), 4.99 (d, *J* = 12.7 Hz, 1H).

¹³C NMR (126 MHz, Acetone-d₆ + 10 μ L H₂O) δ 208.6, 144.6, 139.4, 134.8, 129.1, 123.8, 102.4, 72.5.

¹¹**B** NMR (160 MHz, Acetone- d_6 + 10 µL H₂O) δ 28.93.

IR (cast film, cm⁻¹): 3389, 3295, 3206, 2936, 2885, 1603, 1449, 1382, 1361, 1080, 1038, 1009. **HRMS (ESI-TOF)**: for C₂₅H₃₃N₆O₅ (M + H)⁺: *calcd*.: 179.0521; *found*: 179.0521.



Compound 13 (Figure 3): A small sample of control compound **13** was prepared by acidic hydrolysis in aqueous methanol, followed by oxidative cleavage of the pinacol group. It was isolated in crude form of moderate purity, as a mixture with **2**, however it was sufficient for obtaining the desired chemical shift data (see NMR spectra in Section 8, 500 MHz, DMSO-d6).

3. Molecular Modeling (Figure 2b)

Molecular modeling was performed using Mac Spartan 18. The equilibrium conformation and energy of all compounds below were obtained using the semi-empirical AM1 method (gas phase) [Notes: Energy values are shown in KJ/mol. Optimized energy of H_2O : –247.9 KJ/mol is not included in the energy difference but is constant for all equations]. These calculations support the notion that the open form is much more favorable in the case of compound **2** that benzoxaborole.



4. High Resolution ESI-MS Spectra

4.1 Compound 2



4.2 Benzoxaborole



4.3 Compound 12





5. NMR pKa Titration Data

The pKa of benzoxaborole and compound **2** was determined as follows: ²

A phosphate buffer solution was prepared by dissolving 690 mg NaHPO₃ in 5 mL D₂O and \sim 30 mL H₂O in a 50 mL volumetric flask, then diluting to 50 mL total volume. In a second 25 mL volumetric, 0.5 mmol of the either benzoxaborole or compound **2** was dissolved in a minimum amount of deuterated DMSO, then diluted to 25 mL total volume with the phosphate buffer

solution. A volume of ~1.0 mL of the solution was added to 12 separate vials, and the pH of each was adjusted with NaOH (0.1% and 1%) and HCl (0.3%) so that a range of pH's between ~2 and ~11 was achieved. The various solutions were transferred to NMR tubes and their ¹¹B NMR spectra were recorded. The ¹¹B chemical shift was plotted against the pH and from this the pKa of the boron containing compound was determined. NMR lock was done on D_2O .

The results of the NMR ¹¹B shifts at various pH are shown below:



Benzoxaborole



	Benzoxaborole	Compound 2	
рН	Boron Shift (ppm)	рН	Boron Shift (ppm)
2.31	32.428	2.87	29.350
3.4	32.369	3.87	28.893
4.47	32.236	4.85	28.984
5.4	31.838	5.84	29.126
6.38	29.048	6.73	28.998
7.31	19.178	7.18	28.691
8.56	9.367	7.82	8.393
9.41	8.384	8.27	7.904
10.47	8.146	9.1	7.602
11.27	8.160	10.07	7.643
		11.37	7.574

5.1 Benzoxaborole Acid/Base Titration Curve



6. Amino Group, Amino Acid and Lysozyme Conjugation

Conjugation experiments of compound **2** with amino group containing small molecules were performed in deuterated acetone at a concentration of 0.05 mM in a 1:1 ratio of boronic acid to amino acid. Each reaction was stirred at room temperature for 30 min then immediately submitted for analysis by ¹H NMR spectroscopy. A small aliquot of each reaction was also diluted in ACN and submitted for LC-MS and the data analyzed using Mass Hunter software.

Amino acid conjugation was performed in 50 mM ammonium acetate buffered D_2O at a concentration of 0.05 mM of both compound **2** and either cysteine, ethyl ester *O*-protected cysteine or serine. In the case of cysteine and ethyl ester *O*-protected cysteine conjugation, the solid white material that crashed out of solution was collected by filtration, dried in vacuo, then dissolved in deuterated acetone for ¹H NMR analysis. A small amount of material was also

submitted to HRMS. Serine conjugation produced no such precipitate and was therefore allowed to stir for 24 hours before being analyzed by ¹H NMR.

Lysozyme conjugate studies were performed by mixing 10 μ M lysozyme with 10 mM 2-FPBA, 2-APBA or compound **2** in 2 mL of 50 mM ammonium acetate buffer and allowed to stir at room temperature for 30 mins. A 100 μ L aliquot of each reaction was diluted with 100 μ L ACN and subjected to ESI-FTICR-MS. For fructose competition studies, 10 mM of fructose was added to the 2 mL conjugation mixtures, stirred for 5 mins, followed by another 100 μ L aliquot removed and diluted with 100 μ L ACN and analyzed by ESI-FTICR-MS. Each peak, m/z = (M+8H)⁸⁺, (i.e, lysozyme + number of boronic acid (BA) conjugated + 8H+)/8. For every boronic acid conjugation seen on lysozyme, there is one or two H₂O lost, and a mass peak for each unit. When there are 3 boronic acid conjugations, there can be up to 6 mass peaks for 1 – 6 H₂O lost. For example:





6.1 Compound 2 Lysozyme Conjugation MS (Figure 12)

ESI-FTICR-MS lysozyme binding assays showing the $(M+8H)^{8+}$ charge state of lysozyme with excess boronic acid (BA) **2** (top) and competition experiment with excess fructose (bottom).



6.2 2-FBPA Lysozyme Conjugation MS

ESI-FTICR-MS lysozyme binding assays showing the $(M+8H)^{8+}$ charge state of lysozyme with excess 2-FPBA (BA) **2** (top) and competition experiment with excess fructose (bottom).



6.3 2-APBA Lysozyme Conjugation MS

ESI-FTICR-MS lysozyme binding assays showing the $(M+8H)^{8+}$ charge state of lysozyme with excess 2-APBA (BA) 2 (top) and competition experiment with excess fructose (bottom).



6.4 Compound 2 and L-Cysteine Conjugate: NMR and MS Spectra

NMR spectra of (a) compound **2** (b) ethyl ester L-cysteine and (c) compound **2** conjugate with ethyl ester L-cysteine. Solvent: acetone-d6 with one drop H_2O .



S22



 1 H, 13 C, and 11 B NMR spectra of compound **2** L-cysteine conjugate. Solvent: DMSO-d6 with one drop H₂O.



HRMS of compound 2 L-cysteine conjugate.



6.5 LC-MS of Compound 2 Conjugation with Benzylamine



6.6 LC-MS of Compound 2 Conjugation with 2-Hydroxybenzylamine

7. References

- 1. S. Luliński, J. Serwatowski, J. Organomet. Chem. 2007, 692, 14, 2924.
- 2. Ricardo, C. L.; Mo, X.; McCubbin, J. A.; Hall, D. G. Chem. Eur. J. 2015, 21, 4218.

8. NMR Spectra















