JOC*Note*

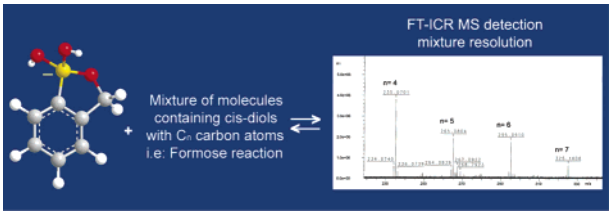
2-Hydroxymethylboronate as a Reagent To Detect Carbohydrates: Application to the Analysis of the Formose Reaction

Alonso Ricardo,[§,||] Fabianne Frye,[†,⊥] Matthew A. Carrigan,[‡] Jeremiah D. Tipton,[†,#] David H. Powell,[†] and Steven A. Benner*[,‡]

Department of Chemistry, University of Florida, Gainesville, Florida 32611, Foundation for Applied Molecular Evolution, Gainesville, Florida, 32601, and Departamento de Química, Universidad de los Andes, Bogota, Colombia

benner@ffame.org; aricardo@molbio.mgh.harvard.edu

Received August 28, 2006



2-Hydroxymethylphenylboronate is described as a reagent that converts neutral 1,2-diols, as found in simple carbohydrates, into 1:1 anionic complexes that are easily detected by Fourier transform ion cyclotron resonance mass spectrometry. The value of this reagent was demonstrated through its application to analyze complex mixtures of carbohydrates formed in the formose process, often cited as a way that biologically significant carbohydrates might have been generated from formaldehyde under prebiotic conditions. Coupled with isotope studies, the reagent shows that the simplest autocatalytic cycle for the consumption of formaldehyde in this process cannot account for the bulk consumption of formaldehyde.

As part of our work seeking to understand how borate minerals might moderate the formose reaction to generate carbohydrates possibly useful in the prebiotic synthesis of RNA,[1] we needed to analyze complex mixtures of carbohydrates. The formose reaction generates these via the conversion of formaldehyde in a calcium hydroxide suspension to (presumably) glycolaldehyde via a reaction whose mechanism is unknown but which is very slow.[2,3] Once formed, the enolizable glycolaldehyde reacts rapidly with more formaldehyde to generate triose products that themselves can act as both nucleophiles and electrophiles in subsequent aldol reactions. The consequence is a cascade of aldol reactions and catalytic cycles (Figure 1, expanded figure in the Supporting Information) that generates a brown tar that has never been fully characterized. As the figures suggest, there are many potential products.

The detection of carbohydrates in aqueous solution is, of course, a major challenge for biological chemistry and in other areas.[4a−g] Several groups have suggested the use of boronic acids for this purpose, on the basis of the ability of boronate to form complexes with the 1,2-diol units. For example, borate itself complexes two 1,2-diol units to form a 2:1 complex that has four boron−oxygen bonds. This complex is anionic and therefore easily visible by electrospray ionization mass spectrometry run in negative ion mode; the parent carbohydrates are not.[5a,b]

For the purpose of mass spectrometric analysis of carbohydrates in complex mixtures, the anionic complex formed by borate itself is problematic, as the two carbohydrates complexed to a single boron atom need not be the same, and the number of complexes scales with the square of the number of carbohydrates in the mixture. We reasoned that a phenylboronate might be a preferred complexing reagent, especially if it contained a 2-hydroxymethyl substituent, as the 2-hydroxymethyl substituent would itself donate the fourth bond to boron (Figure 2). This is reminiscent of the 2-dialkylaminophenylboronates discussed in 1982 by Wulff as reagents to capture carbohydrates.[6] Further, after our work was completed, Dowlut and Hall proposed 2-hydroxymethylphenylboronate, available from ASDI Inc. (Newark, DE), for the analysis of simpler mixtures of carbohydrates using NMR. They measured the binding constants of this reagent to several biologically important carbohydrates.[7]

To illustrate the application of this reagent to analyze complex mixtures arising from the formose reaction, a solution containing formaldehyde, $CaCl_2$, and sodium hydroxide was incubated at 65 °C. Aliquots from this mixture were then treated with a solution of 2-hydroxymethylphenylboronic acid, and the anionic complexes were analyzed by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). Carbohydrates were assigned by identifying peaks having masses of $117 + (n \times 30 - 2)$; this formula includes the mass of the reagent core with the mass of the carbohydrate, with n indicating the number of carbons in the carbohydrate, minus the two protons removed when the complex is formed. The isotopic signature character-

§ Departamento de Química, Universidad de los Andes.
|| Current address: Simches Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114.
† Department of Chemistry, University of Florida.
⊥ Current address: Department of Chemistry, University of California Irvine, Irvine, CA 92697.
‡ Foundation for Applied Molecular Evolution.
# Current address: The Scripps Research Institute, FL.

(1) (a) Ricardo, A.; Carrigan, M. A.; Olcott, A. N.; Benner, S. A. *Science* **2004**, *303*, 196. (b) Prieur, B. *Compt. Rend. Acad. Sci.* (Paris, Chimie) **2001**, *4*, 667−670.
(2) Butlerov, A. *C. R. Acad. Sci.* **1861**, *53*, 145−147.
(3) Breslow, R. *Tetrahedron Lett.* **1959**, *21*, 22−26.
(4) (a) Davis, A. P.; Wareham, R. S. *Angew. Chem., Int. Ed.* **1999**, *38*, 2978−2996. (b) Rohrscheib, M.; Robinson, R.; Eaton, R. P. *Diabetes, Obes. Metab.* **2003**, *5*, 280−284. (c) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chem., Int. Ed.* **1996**, *35*, 1910−1922. (d) James, T. D.; Shinkai, S. *Top. Curr. Chem.* **2002**, *218*, 159−200. (e) Wang, W.; Gao, X.; Wang, B. *Curr. Org. Chem.* **2002**, *6*, 1285−1317. (f) Wiskur, S. L.; Ait-Haddou, H.; Lavigne, J. J.; Anslyn, E. V. *Acc. Chem. Res.* **2001**, *34*, 963−972. (g) Wang, B.; Yang, W.; Fan, H.; Gao, X.; Gao, S.; Karnati, V. V. R.; Ni, W.; Hooks, W. B.; Carson, J.; Weston, B.; Wang, B. *Chem. Biol.* **2004**, *11*, 439−448.
(5) (a) Qian, L.; Ricardo, A.; Winefordner, J.; Smith, B.; Benner, S. A.; Powell, D. H. *Anal. Chem.* **2005**, *77*, 4503−4508. (b) Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349−351.
(6) Wulff, G. *Pure Appl. Chem.* **1982**, *54*, 2093−2102.
(7) Dowlut, M.; Hall, D. G. *J. Am. Chem. Soc.* **2006**, *128*, 4226−4227.

10.1021/jo061770h CCC: $33.50
Published on Web 11/16/2006

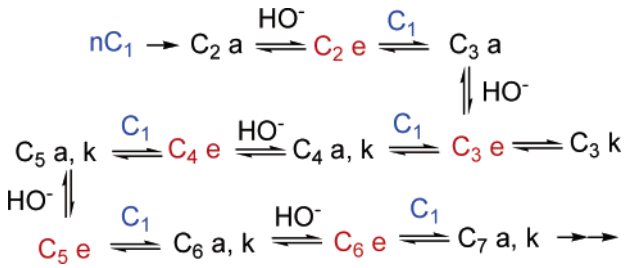


FIGURE 1. In the formose reaction, C_1 (formaldehyde) reacting with another formaldehyde molecule to yield a C_2 compound (glycolaldehyde), which then enolizes and undergoes further aldol additions to yield the higher carbohydrates (abbreviations specify the number of carbon atoms). This process continues until branching, a Cannizarro reaction, or another process removes acidic protons adjacent to a carbonyl group. In autocatalytic cycles, C_4, C_5, and higher species may suffer retroaldol reactions to yield C_2 and C_3 species, which re-enter the cycle and consume formaldehyde. a = aldehyde; k = ketone; e = enediol. For a more detailed scheme, see the Supporting Information.

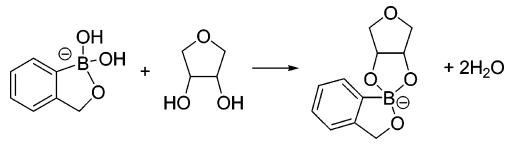


FIGURE 2. Reaction in water between 2-hydroxymethylboronate and a neutral *cis*-1,2-diol yielding an anionic complex detectable by FT−ICR MS. 1,4-Anhydroerythritol is shown as an example.

istic of boron was also key, in particular, to recognize peaks arising through the complexation of two borons.

This allowed us to follow the fate of the carbon in the formose process as a function of time under a variety of conditions. Figure 3a shows a typical MS spectra for the reaction. The remaining panels of Figure 3 show the results obtained at 68 °C (where the mixture turns yellow at 18 min) and 40 °C (where the mixture turns yellow at 135 min) for the total detectable carbon in the reaction. Independent measurement of the concentration of formaldehyde using the chromotropic acid assay[8] showed that the concentration of formaldehyde decreased dramatically at the yellowing point (data not shown).

Our results complement those of Dowlut and Hall,[7] who did not study the formose process. Their results were more quantitative and showed the utility of 2-hydroxymethylphenylboronate to analyze carbohydrates using mass spectrometry. Mass spectrometry of the 2-hydroxymethylphenylboronate complexes cannot, of course, distinguish between isomeric forms of the carbohydrates being presented. Further, quantitation between the species is only approximate, as different diols have different dissociation constants, and the ion response need not be identical for all species (although these ions have very similar structures overall). Even with these limitations, these traces are valuable, as they provide several insights into the formose process that have been elusive for a century and are especially interesting to the prebiotic synthesis community.

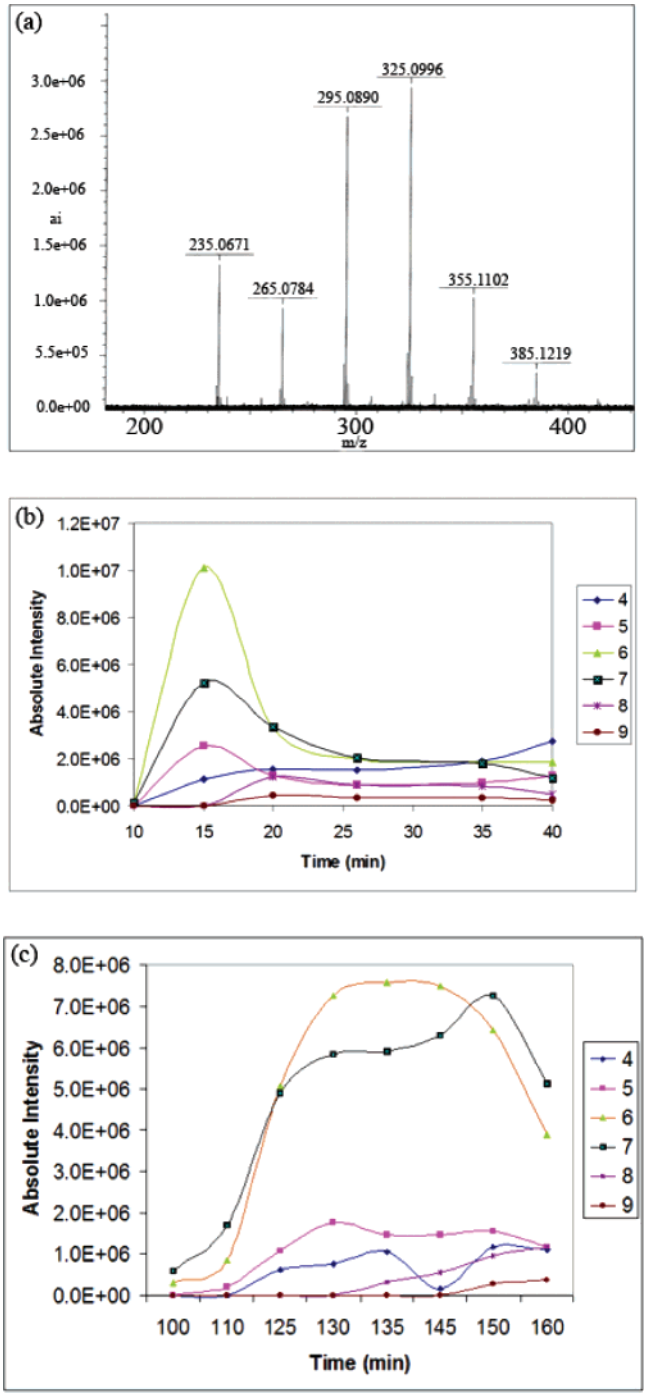


FIGURE 3. (a) Representative trace of FT−ICR analysis of formose mixture. Signals for species having between four (C4) and nine (C9) carbon atoms are seen. (b) Progress curve of the formose reaction at 68 °C. The lines show, for species having the indicated number of carbon atoms, the change in time of the amount of each carbon species detected in the experiment, yellowing occurs at 18 min. (c) As before, but at 40 °C, yellowing occurs at 145 min.

First, no C3 carbohydrates were detected after extended incubation. Although these might be missed due to weak binding, the fact that they are not observed even at high concentrations of the boronate reagent suggests that they do not accumulate in large amounts. The higher C4, C5, and C6 species are, however, easily detected by the reagent before yellowing. Thus, C6 (*m*/*e* of the complex = 295.098) and C7

(8) Altshuller, A. P.; Miller, D. L.; Sleva, S. F. *Anal. Chem.* **1962**, *34*, 621−625.

(9) (a) Shigemasa, Y.; Sakai, H.; Nakashima, R. *Nippon Kagaku Kaishi* **1982**, *10*, 1626−1632. (b) Shigemasa, Y.; Akagi, S.; Nakashima, R.; Saito, S. *Carbohydr. Res.* **1980**, *80*, C1−3. (c) Shigemasa, Y.; Akagi, S.; Shinichi, W. E., Nakashima, R. *J. Catal.* **1981**, *69*, 58−68. (d) Shigemasa, Y.; Nagae, O.; Sakazawa, C.; Nakashima, R. *J. Am. Chem. Soc.* **1978**, *100*, 1309−1310.

(10) Huskey, W. P.; Epstein, I. R. *J. Am. Chem. Soc.* **1989**, *111*, 3157−3163.

(*m*/*e* of the complex = 325.109) carbohydrates dominate the mixture at times approaching the yellowing point. The observation of the C7 species is particularly important, as one C7 species (1,2,4,5-tetrahydroxy-2,4-dihydroxymethylpentan-3-one) was isolated as a major component from a formose mixture by Shigemasa et al.[9a,b] and analyzed by ^{13}C and ^{1}H NMR and periodate degradation. We repeated the synthesis and found it reproducible in every detail and consistent with our mass spectrometry results.[9c,d] The dominance of C6 and C7 species late in the pre-yellowing formose process was observed with the 2-hydroxymethylphenylboronate reagent at all of the conditions examined. This suggests that the quantitation, although not precise, is not wrong by orders of magnitude. Further, after the yellowing was well under way, a mixture of C6, C7, C8, and C9 carbohydrates is seen in an approximate ratio of 1:2.5:1:0.3.

The reagent also allowed us to rule out the simplest autocatalytic cycle for the consumption of formaldehyde, which involves the conversion of glyceraldehyde to a tetrulose, the isomerization of the tetrulose to a tetrose, followed by its retroaldol cleavage to form two glycolaldehyde molecules (see Supporting Information). To do this, the reaction was repeated at 65 °C in D_2O and *in situ* generated $Ca(OD)_2$ as the base. Under these conditions, yellowing began at 39 min, indicating a solvent deuterium kinetic isotope effect (compare to the report of Huskey and Epstein).[10] Nevertheless, after 32 and 34 min, the peak at 326, corresponding to the C7 sugar complex that incorporated a single deuterium, was only about 10% of the whole (after correction for ^{13}C natural abundance). This shows that, before yellowing, few of the intermediate enols protonate, and very few retroaldol reactions expel formaldehyde.

This rules out the simplest autocatalytic cycle, where the tetrulose is converted to the isomeric tetrose, which suffers a retroaldol reaction to give two glycolaldehydes. This would require the introduction of two deuteriums from solvent, necessarily generating deuterated C6 and C7 species. These were not observed.

Experimental Section

Reagents and Materials. 2-Hydroxymethylphenylboronic acid was purchased from a commercail supplier; sodium deuteroxide was purchased as a solution (10 M, D_2O) from a commercial supplier. All other reagents were purchased from a commercial supplier in their highest purity.

Analysis of Carbohydrates Formed during the Formose Process. In a typical reaction, a solution containing formaldehyde (350 mM, generated from paraformaldehyde), $CaCl_2$ (50 mM), and sodium hydroxide (60 mM) was incubated at 65 °C. From this, aliquots (0.50 mL, every 2 min) were removed and added to a solution of 2-hydroxymethylphenylboronic acid (0.20 mL, 0.1 M, pH adjusted to 9−10 with NaOH), followed by acidification with Amberlite resin (CG-50 H^+ form). After acidification, a portion of the mixture (0.40 mL) was removed and added to 4-methyl morpholine buffer (0.40 mL, 50 mM, pH 8.7). The samples were diluted 10-fold in Milli-Q water, mixed 1:1 with acetonitrile, and injected into the FT-ICR instrument.

Deuterium incorporation experiments were done as described above, except that deuterium oxide was used instead of water, and sodium deuteroxide was used instead of the hydroxide.

Instrumentation. All FT-ICR analysis were done in a Bruker APEX II FT−ICR MS. Samples were diluted 10-fold in Milli-Q water and mixed 1:1 with acetonitrile before injection into the instrument.

Acknowledgment. We are grateful to the NASA Exobiology Program for support of this work and dedicate this work to the memory of Lidia N. Matveeva.

Supporting Information Available: Mass spectrometry data, additional figures, and a detailed mechanistic scheme. This material is available free of charge via the Internet at http://pubs.acs.org.

JO061770H