Phytochemistry, Vol. 21, No. 9, pp. 2301–2303, 1982.
Printed in Great Britain.

0031-9422/82/092301-03$03.00/0

HYDROGEN FLUORIDE SACCHARIFICATION OF CELLULOSE AND XYLAN: ISOLATION OF α-D-GLUCOPYRANOSYL FLUORIDE and α-D-XYLOPYRANOSYL FLUORIDE INTERMEDIATES, AND 1, 6-ANHYDRO-β-D-GLUCOPYRANOSE

HAIM HARDT* and DEREK T. A. LAMPORT

MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824, U.S.A.

(*Received* 16 *February* 1982)

Key Word Index—Hydrogen fluoride saccharification; cellulose; xylan; α-glycosyl fluorides.

Abstract—Termination of the hydrogen fluoride solvolysis of cellulose and of xylan by instantaneous neutralization enabled the identification of the solvolysis intermediates exclusively as the α-glycosyl fluoride anomers. A 1, 6-anhydroglucose was identified as a minor degradation product from cellulose.

Recently we began a re-evaluation of hydrogen fluoride (HF) solvolysis of wood [1] as a possible alternative saccharification process [2, 3]. As a first step we used pure cellulose as a model compound and to prevent the formation of oligosaccharides through acid-catalysed reversion we terminated the solvolysis by rapid neutralization with calcium carbonate. Consequently, we were able to isolate and identify the reaction intermediate between cellulose and reversion oligosaccharides as α-D-glucopyranosyl fluoride (**1**) [3].

After cellulose, xylan is the second major polysaccharide in hardwoods accounting for 15–25% of the total wood carbohydrates. Xylan, like cellulose, consists of β-1, 4-linked linear chains. The present report describes the isolation of the intermediate α-D-xylopyranosyl fluoride (**3**) from the solvolysis of purified xylan by similar procedures and also the identification of a minor degradation product obtained from cellulose, as 1,6-anhydro-β-D-glucopyranose (**2**).

The reaction products obtained after HF solvolysis of cellulose followed by rapid calcium carbonate neutralization were separated by PC. The area suspected of containing **1** as well as that which contained **2** were each eluted in a small volume of water. The GC of the TMSi derivatives of the water soluble products before purification by PC showed a major new peak, assigned to TMSi-glucosyl fluoride, with R_t of 18.46 min, and a RR_t of 0.86 and a minor new component which corresponded, as shown below, to **2**, with R_t of 14.29 min and RR_t 0.66.

The GC/MS analysis of **4**, the TMSi derivative of **1**, was done after purification by PC. The M^+ at *m/z* 470 matched the MW of the assumed silylated compound and in addition, the major fragments at *m/z* 450 $[M-HF]^+$ and at 435 $[M-Me-HF]^+$ established the identification of the 18.46 min GC component as 2,3,-4,6-tetra-*O*-trimethylsilyl-D-glucopyranosyl fluoride. The specific rotation of the glucosyl fluoride indicated the presence of the α-anomer. Furthermore, of significance was the occurrence of a single GC peak of glucosyl fluoride which indicated that it was obtained exclusively in one anomeric form. It is well known that the α-anomer of any glucopyranosyl halide is the more stable form whereas the less stable β-anomer isomerizes rapidly to the α-form, even if the β-anomer is initially formed (the anomeric effect [4, 5]). On this basis, and in view of our optical rotation and TLC results [3], we assigned the TMSi-glucosyl fluoride product to the α-anomeric form **4**.

If the splitting of glycosidic bonds in anhydrous HF produced, at least initially, both α- and β-glycosyl fluoride, one must deduce that the mechanism of this nucleophilic substitution occurs via an SN_1 mechanism. The first step in such a mechanism is the protonation of the glycosidic oxygen and the production of the HF_2^- anion [6], followed by the formation of a carbocation intermediate (referred to as glycosyl cation [7]). This intermediate reacts finally with HF_2^- producing both the α- and β-anomers of glycosyl fluoride.

The strong anomeric effect exhibited by fluorine probably excludes an alternative and speculative hypothesis for the preferred production of the α-anomer which was to assume a nucleophilic attack by the HF fluorine on the β-oriented and rigidly held C-1 glycosidic bonds in the cellulose, via a concerted single step SN_2 mechanism. Still, to test this hypothesis we replaced cellulose with amylose in which the glycosidic bonds are oriented at C-1 and expected to observe the β-glucosyl fluoride intermediate as predicted from such a mechanism. The observed GC peaks, however, were again only that of **4** corresponding to the α-anomer, and that of **5** the TMSi derivative of **2**. The absence of a new GC peak

*Present address: Department of Biochemistry, The University of Connecticut Health Center, Farmington, CT 06032, U.S.A.

corresponding to a β-anomer was evident even when the TMSi derivatization of the sugar products was done immediately after the solvolysis termination, e.g. on samples from the sugar-containing calcium carbonate–calcium fluoride precipitate. The best yield of **1** from HF solvolysis of cellulose was 20.4%, on a weight basis. α-Glycosyl fluorides are traditionally synthesized by a short treatment of the acylated aldoses with HF followed by deacetylation (cf. refs. [8, 9]).

As mentioned above, component **2** whose TMSi derivative **5** appeared at R_t 14.29 min and RR_t 0.66, was also separated by PC but contained virtually no fluoride. The GC/MS of **5** showed a M^+ at 378 and a fragment at 363. However, both these fragments were detectable only with a low ionization voltage of 20 eV whereas at 70 eV the intensive characteristic fragment at *m/z* 333 was always detected as the highest mass. Otherwise our pattern of fragmentation agreed well with that published by Hyns and Scharmann [10] for **5** except for the M^+ which they could not detect, using 70 eV. The identification of the 14.29 min GC component as **5** was eventually confirmed by GC of the TMSi derivative of commercial **2** (Sigma lot 70F-0686). Compound **2**, known also as 'levoglucosan', was reported by Helferich and Böttger [11] as a component of the 'cellan' produced by the reaction of HF with cellulose. In spite of these results it was still possible that compound **2** was formed by an alkaline catalysed reaction from **1** (cf. ref. [9]) during the neutralization with calcium carbonate and not by the acid catalysed reactions in HF. To exclude this possibility we terminated the HF solvolysis of cellulose by a 1.5 hr evacuation rather than by neutralization. The GC analysis of the water-soluble sugars as their TMSi derivatives, however, again showed the presence of **5** and the exclusive formation of the α-anomer **4**. These results, therefore, also excluded a possible specific destruction of β-glucosyl fluoride by the neutralization procedure. TMSi derivatization by itself has been shown not to affect the anomeric configuration of monosaccharides [12].

HF solvolysis of xylan was terminated with calcium carbonate (see Experimental) and the water-soluble sugars were analysed by GC of their TMSi derivatives before and after separation by PC. Before PC, the residual glucan in the xylan yielded peaks at 19.84 and 21.28 min corresponding to α- and β-D-glucose and the 18.20 min peak of **4**. The R_t of the TMSi-mannitol standard was 21.28 min. The xylan yielded the peaks corresponding to α- and β-D-xylose at 14.76 and 16.31 min and a new peak at R_t 11.05 min and RR_t 0.52 which we attributed to **6** the TMSi derivative of α-D-xylopyranosyl fluoride **3**. After separation by PC the band containing **3** was at R_f 0.51 and was eluted in a small volume of water and analysed by GC of TMSi derivatives which showed that the putative **3** was obtained virtually free of sugar contaminants. For the determination of the fluoride–xylose molar ratio, we used samples from the eluted band. In one sample we assayed xylose via GC of the TMSi-*O*-methyl glycosides. In another sample we measured the concentration of free fluoride anion before and after hydrolysis with 180 mN sulfuric acid at 95° for 15 min [13]. Free fluoride in the unhydrolysed sample was negligible. The molar ratio xylose–fluoride in the hydrolysed sample was 1:1 indicating the presence of **3**. TLC analysis yielded a single spot at R_f 0.58–0.71 compared to the xylose marker spot at R_f 0.38.

GC/MS analysis confirmed the identification of xylosyl fluoride as its TMSi derivative. The M^+ at *m/z* 368 matched the MW of the silylated xylosyl fluoride and the major fragments at *m/z* 348 $[M-HF]^+$ and at 333 $[M-Me-HF]^+$ established the identification of the 11.05 min GC component as 2, 3, 4 tri-*O*-trimethylsilyl-D-xylopyranosyl fluoride. We determined the relative molar response factor of the putative xylosyl fluoride by assuming that its molar concentration was equal to that of the xylose determined after methanolysis, and GC of the TMSi derivatives. The molar response factor of TMSi-xylosyl fluoride relative to TMSi-mannitol was 2.3 and using this value we calculated that the yield of xylosyl fluoride was 30 mg/g of xylan, i.e. 3%.

Specific rotation measurements using the aqueous solution of the purified xylosyl fluoride gave the value $[\alpha]_D^{24} = +60°$ (H_2O; *c* 0.05) which is similar to the lit. values $[\alpha]_D^{24} = +55° \pm 10$ (H_2O; *c* 0.2) [14] or $[\alpha]_D^{20} = +73° \pm 3$ (H_2O; *c* 0.03) [15]. We therefore assigned the xylosyl fluoride product to the α-anomeric form **3**. As in the case of **1** the xylosyl fluoride was obtained exclusively in one anomeric form which is again explained by the strong anomeric effect of fluorine, i.e. by the favorable production of the more stable α-anomer. Unlike **1**, in which the hydroxyl group at C-6 is favorably placed for intramolecular attack leading to the substitution of the fluorine at C-1 and to the formation of the 1,6-anhydrobridge, **3** was not expected to yield any stable intramolecular bond and therefore gave the free sugar as the only primary degradation product. This different behavior of C_6- and C_5-glycopyranosyl fluorides was observed earlier by Barnett under alkaline hydrolysis conditions [15].

EXPERIMENTAL

Reaction of cellulosic materials with HF was contained in a Kel-F vacuum distillation system [16], (Peninsula Laboratories Inc.). Liquid MF was purchased from Matheson Gas Products Co. Filter paper, Whatman No. 4, which consists of 99% cellulose was used as cellulose standard. Larchwood xylan which was purchased from Sigma (lot 62-C-2820) required further purification from residual mannan and glucan. After purification by pptation at pH 4.5 [17] the recrystallized xylan still contained *ca* 9% residual glucan and 2% residual mannan on a molar basis as shown by GC of its alditol acetate derivatives.

Typical HF solvolysis expts of dry polysaccharides were initiated by cooling the reaction container with liquid N_2 and evacuating it for 10 min at *ca* 10 mm Hg. Afterwards we distilled *ca* 10 ml dry HF (dried by storage over CoF_3 in a reservoir) over the 1 g xylan sample. The frozen sample and HF were then allowed to warm up and the reaction was finally run at 0° or at room temp. until a clear soln was obtained (*ca* 20 min). For termination, the liquid HF soln was poured into a polypropylene beaker containing a suspension of 30 g $CaCO_3$ in 150 ml CH_2Cl_2 cooled in liquid N_2 [18]. After separating and discarding the CH_2Cl_2, the H_2O-soluble sugars were extracted from the solid cake in a large vol. of H_2O, freeze-dried and redissolved in a small vol. of H_2O. Samples from such preparations were further separated by descending PC on Whatman No. 1, No. 4 or 3 MM filter paper for 12–48 hr using as eluant the epiphase

of a *n*-BuOH–EtOH–H_2O (4:1:5) mixture and staining with alkaline $AgNO_3$ [19]. TLC was on Si gel G plates (500 μm thickness) with the same eluant used for PC.

The fluoride concn of aq. solns was determined with an Orion fluoride selective combination electrode attached to a Corning pH/ion meter. GC with FID was via 72 in. × 1.5 mm columns with 3% SP 2100 as stationary phase and He carrier at a flow rate of 40 ml/min. We analysed sugars by GC either as their TMSi derivatives or as the TMSi-*O*-methyl glucosides [20] using mannitol as int. standard, or as their alditol acetates [21] using inositol as int. standard. The second method involved a methanolysis step in MeOH–1.5 N HCl at 95° for 90 min. Running conditions, after an initial delay of 4 min, were for the TMSi derivatives: 120–225° at 4°/min, and for TMSi-*O*-methyl derivatives: 120–180° at .1°/min. For GC of alditol acetates we used 1% OV-275 as stationary phase and the running conditions were: 130–230° at 1.5°/min and an initial delay of 2 min.

GC/MS analyses were performed using electron impact mode with an ionizing voltage of 70 or 20 eV.

α-D-Glucopyranosyl fluoride (**1**). PC on Whatman No. 4, R_f 0.38 (BuOH–EtOH–H_2O, 4:1:5); $[\alpha]_D^{24}$ + 106° (H_2O; *c* 0.5). 2,3,4,6-*Tetra*-O-*trimethylsilyl-α-D-glucopyranosyl fluoride* (**4**). The TMSi derivative of **1**. GC RR_t 0.86; GC/MS at 70 eV *m/z* (%): 470 $[M]^+$ (2), 455 $[M-Me]^+$ (2), 450 $[M-HF]^+$ (4), 435 $[M-Me-HF]^+$ (6), 365 (3), 345 (4), 331 (11), 319 (11), 305 (5), 291 (10), 271 (3), 257 (8), 243 (13), 217 (100), 204 (40).

1, 6-*Anhydro-β-D-glucopyranose* (**2**). PC on Whatman No. 4, R_f 0.48; on Whatman 3 MM, R_f 0.40 (BuOH–EtOH–H_2O, 4:1:5). 2,3,4-*Tri*-O-*trimethylsilyl*-1,6-*anhydro-β-D-glucopyranose* (**5**). The TMSi derivative of **2**. GC RR_t 0.66; GC/MS at 20 eV *m/z* (%): 378 $[M]^+$ (0.26); 363 $[M-Me]^+$ (0.59), 349 (0.45), 335 (13), 334 (29), 333 (100), 317 (10), 305 (5), 273 (5), 261 (6), 260 (9), 243 (23), 217 (247), 204 (266).

α-D-Xylopyranosyl fluoride (**3**). PC on Whatman No. 4, R_f 0.51 (BuOH–EtOH–H_2O, 4:1:5); TLC (Si gel G; BuOH–EtOH–H_2O, 4:1:5) R_f 0.58–0.71; $[\alpha]_D^{24}$ + 60° (H_2O; *c* 0.05). 2,3,4-*Tri*-O-*trimethylsilyl-α-D-xylopyranosyl fluoride* (**6**). The TMSi derivative of **3**. GC RR_t 0.52; GC/MS at 70 eV *m/z* (%): 368 $[M]^+$ (4), 353 $[M-Me]^+$ (2), 348 $[M-HF]^+$ (5), 333 $[M-Me-HF]^+$ (5), 309 (8), 307 (5), 305 (3), 243 (3), 217 (100), 204 (30).

Acknowledgement—This work was supported by the DOE Contract No. DE-AC02-76ER01338.

REFERENCES

1. Fredenhagen, K. and Cadenbach, G., (1933) *Angew. Chem.* **46**, 113.
2. Lamport, D. T. A., Hardt, H., Smith, G., Mohrlok, S., Hawley, M. C., Chapman, R. and Selke, S. (1981) in *Energy from Biomass, 1st E.C. Conference* (Palz, W., Chartier, P. and Hall, D. O., eds.) p. 292. Applied Science, London.
3. Hardt, H. and Lamport, D. T. A. (1982) *Biotechnol. Bioeng.* (in press).
4. Edward, J. T. (1955) *Chem. Ind. (London)* 1102.
5. Lemieux, R. V. and Chu, N. J. (1958) 133*rd Meeting Am. Chem. Soc., Abstracts of Papers*, 31N.
6. Kilpatrick, M. and Jones, J. G. (1967) in *Chemistry of Non Aqueous Solvents* (Lagowski M., ed.) Vol. 2, p. 43. Academic Press, New York.
7. Bochkov, A. F. and Zaikov, G. E. (1979) *Chemistry of the O-Glycosidic Bond*, Chapter 6, p. 181. Pergamon Press, Oxford.
8. Penglis, A. A. E. (1981) in *Advances in Carbohydrate Chemistry and Biochemistry* (Tipson, S. S. and Horton, D., eds.) Vol. 38, p. 195. Academic Press, New York.
9. Micheel F. and Klemer, A. (1961) in *Advances in Carbohydrate Chemistry* (Wolfrom, M. L. and Tipson, R. S., eds.) Vol. 16, p. 85. Academic Press, New York.
10. Heyns, K. and Scharmann, H. (1966) *Chem. Ber.* **99**, 3461.
11. Helferech, B. and Böttger, A. (1929) *Ann.* **476**, 150.
12. Sweeley, C. C., Bentley, R., Makita, M. and Wells, W. W. (1963) *J. Am. Chem. Soc.* **85**, 2497.
13. Figures, W. R. and Edwards, J. R. (1976) *Carbohydr. Res.* **48**, 245.
14. Barnett, J. E. G., Jarvis, W. T. S. and Munday, K. A. (1967) *Biochem. J.* **105**, 669.
15. Barnett, J. E. (1969) *Carbohydr. Res.* **9**, 21.
16. Sakakibara, S., Shimonishi, Y., Kishida, Y., Okada, M. and Sugihara, H. (1967) *Bull. Chem. Soc. Jpn.* **40**, 2164.
17. Adams, G. A. (1965) in *Methods in Carbohydrate Chemistry* (Whistler, R. L., ed.) Vol. 5, p. 170. Academic Press, New York.
18. Mort, A. and Parker, S. in *Biotechnology R & D for Fuels and Chemicals for Biomass* (Villet, R., ed.) IPC, London (in press).
19. Trevelyan, W. E., Procter, D. P. and Harrison, J. S. (1950) *Nature (London)* **166**, 444.
20. Bhatti, T., Chambers, R. E. and Clamp, J. R. (1970) *Biochim. Biophys. Acta* **222**, 339.
21. Albersheim, P., Nevins, D. J., English, P. D. and Karr, A. (1967) *Carbohydr. Res.* **5**, 340.