ELSEVIER

Carbohydrate Research 276 (1995) 309–320

CARBOHYDRATE RESEARCH

The formation of 2-furaldehyde and formic acid from pentoses in slightly acidic deuterium oxide studied by ^{1}H NMR spectroscopy

Tania Ahmad, Lennart Kenne *, Kjell Olsson, Olof Theander

Department of Chemistry, Swedish University of Agricultural Sciences, Box 7015, S-750 07 Uppsala, Sweden

Received 2 December 1994; accepted 30 March 1995

Abstract

The title reaction at 96°C and pD 1.5, 3.0, or 4.5 was followed by ^{1}H NMR spectroscopy. The rate of pentose degradation increased in the order: arabinose ≈ xylose < ribose < 2-pentuloses. At pD 1.5, the rate of 2-furaldehyde formation increased in the same order. Increasing pD strongly accelerated the degradation of the aldoses but slightly retarded that of the ketoses. Increasing pD also retarded the formation of 2-furaldehyde, particularly from the ketoses, and increased its deuterium content at H-α (from 8–25 to 50–83 atom %) and H-3 (from 79–100 to 100 atom %). This is explained by assuming that 2-furaldehyde had formed mainly via acyclic intermediates, with reversible formation of a 3-deoxypentosulose. The formation of formic acid was slow and did not proceed via 2-furaldehyde. As evident from experiments with 1- or 5-^{13}C-substituted aldopentoses, the formic acid was derived exclusively from the terminal pentose carbons, C-1 being somewhat more important than C-5.

Keywords: Formic acid; 2-Furaldehyde; Pentose

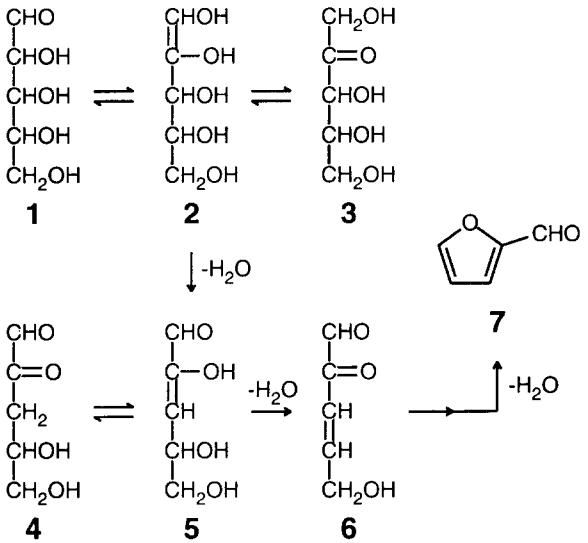
1. Introduction

Dehydration reactions of carbohydrates are important in pulping [1], in cooking (Maillard reaction) [2], and in the manufacture of certain bulk chemicals. Thus, 2-furaldehyde is obtained by the acid-catalyzed dehydration of pentoses and hexuronic acids, present in the polysaccharides of various agricultural waste products [3]. The early literature on such reactions in aqueous media has been reviewed [4].

The first generally accepted reaction route [4] from an aldopentose (**1**) or a ketopen-

* Corresponding author.

SSDI 0008-6215(95)00176-X



Scheme 1. Classical route from pentoses to 2-furaldehyde via acyclic intermediates.

tose (**3**) to 2-furaldehyde (**7**) is shown in Scheme 1 and proceeds via the intermediates **2**, **5**, and **6**. Until fairly recently, sugars were believed to react exclusively through their acyclic forms. The more abundant cyclic forms of **1** and **3** are therefore ignored in Scheme 1. For further simplicity, the stereochemistry is also neglected.

When D-xylose was dehydrated in tritiated 3 M sulfuric acid, the main product **7** contained very little tritium [5]. This was rationalized by assuming that the enols **2** and **5** are dehydrated much faster than they tautomerize. An intramolecular hydride transfer from C-2 to C-1 in the xylose was proposed [6] as an alternative explanation for the absence of tritium at C-α of **7**. More recently, an entirely new reaction route from D-xylose to **7** in strong aqueous acid was suggested [7]. This route involves only cyclic intermediates. The formation of formic acid along with **7** was observed but apparently not explained [7].

In weaker aqueous acid, D-xylose has been degraded into catechols and other compounds, **7** being only a minor product [8]. In aqueous alkali, no **7** was formed, and phenolic products predominated [9]. We have now reinvestigated the transformation of D-xylose and other pentoses into **7** and formic acid in slightly acidic deuterium oxide, using ^{1}H NMR spectroscopy.

2. Results

Three 0.5 M sodium phosphate buffers in deuterium oxide were prepared and adjusted to pD 1.5, 3.0, and 4.5 [1]. Each buffer contained 0.1% (w/v) of pivalic acid, to

[1] pD = Observed pH meter reading + 0.4 [10]; pD also equalled pH of a similar solution of the undeuterated solutes in ordinary water.

be used as an internal ^{1}H NMR standard. The degradation of the following D-pentoses was investigated: arabinose, ribose, ribulose (*erythro*-2-pentulose), xylose, and xylulose (*threo*-2-pentulose). A 2.5% (w/v) solution of each pentose in each buffer was heated at 96°C in a sealed NMR tube.

The ^{1}H NMR spectrum of the solution was recorded at certain times (after transient cooling to 25°C), as exemplified in Fig. 1 for ribose at initial pD 4.5. In each spectrum, the pentose, 2-furaldehyde (**7**), and the formic acid ($^1HCO_2{}^2H$) with intact formyl proton were quantified. The results are shown in Figs 2 and 3.

The isomeric composition of aldopentoses [11] and ketopentoses [12] in deuterium oxide solution at equilibrium has been determined previously, and the ^{1}H NMR spectra of their major isomeric forms have been analyzed [13]. This was very helpful in the interpretation of our spectra. For all pentose signals, our chemical shifts (δ) were 0.02–0.03 ppm lower than those reported [13,14], probably owing to the buffer. In each spectrum, except for that of ribulose, it was possible to select two or more regions which were mutually well resolved and unaffected by hydrogen exchange with the solvent or by overlapping signals from the solvent (HDO) or degradation products. Each of these regions was integrated. Their assignments and positions are listed in Table 1. During each experiment, the ratios of the integrals remained constant within $\pm 5\%$. The pentose concentration was therefore calculated from the average of the integrals.

In **7**, H-4 (δ 7.91) and H-5 (δ 6.75) were integrated; both gave virtually the same results. The formyl proton (δ 8.33) was used in formic acid, and the methyl groups (δ 1.19) in the internal standard, pivalic acid.

As shown in Fig. 2, all three aldoses decomposed significantly, ribose being the most sensitive. Only traces of ketopentoses were observed in the spectra. This could be due to their high degradation rate under the same conditions; see Fig. 3. As shown by Figs 2 and 3, increasing pD strongly accelerated the degradation of the aldoses but slightly retarded that of the ketoses. At pD 1.5, the formation of **7** was fast from the ketoses, slower from ribose, and even slower from the other aldoses. Increasing pD retarded the formation of **7**, notably from the ketoses.

No correction for any decomposition of **7** has been made in Figs 2 and 3. When **7** was heated in phosphate buffers in the same way as the aldopentoses, $< 10\%$ of **7** disappeared; see Fig. 4. In the pentose degradation experiments, additional **7** might have disappeared through reaction with other degradation products. However, the yields of **7**, calculated on consumed pentose and shown in Fig. 5, indicated that very little **7** reacted further. Thus, nearly quantitative such yields of **7** were obtained from the aldopentoses after 72 h at pD 1.5, although Fig. 4 indicates maximum decomposition (9%) of **7** under these conditions. With increasing pD, these yields of **7** dropped quickly (Fig. 5), owing to the formation of other products [8]. These were indicated by strong colour of the solution and by a dark precipitate.

In the aldoses, virtually no hydrogen exchange with the solvent occurred in any of the experiments. In the ketoses, only H-1a and H-1b were exchanged; at pD 3.0 or 4.5, their signals had disappeared after 32 h.

As seen from the signals at δ 9.49 and 7.58 in Fig. 1, part of H-α and most of H-3 in **7** had been replaced by deuterium. The deuterium contents of H-α and H-3 were calculated from the relative decrease of these signals. The results are listed in Table 2.

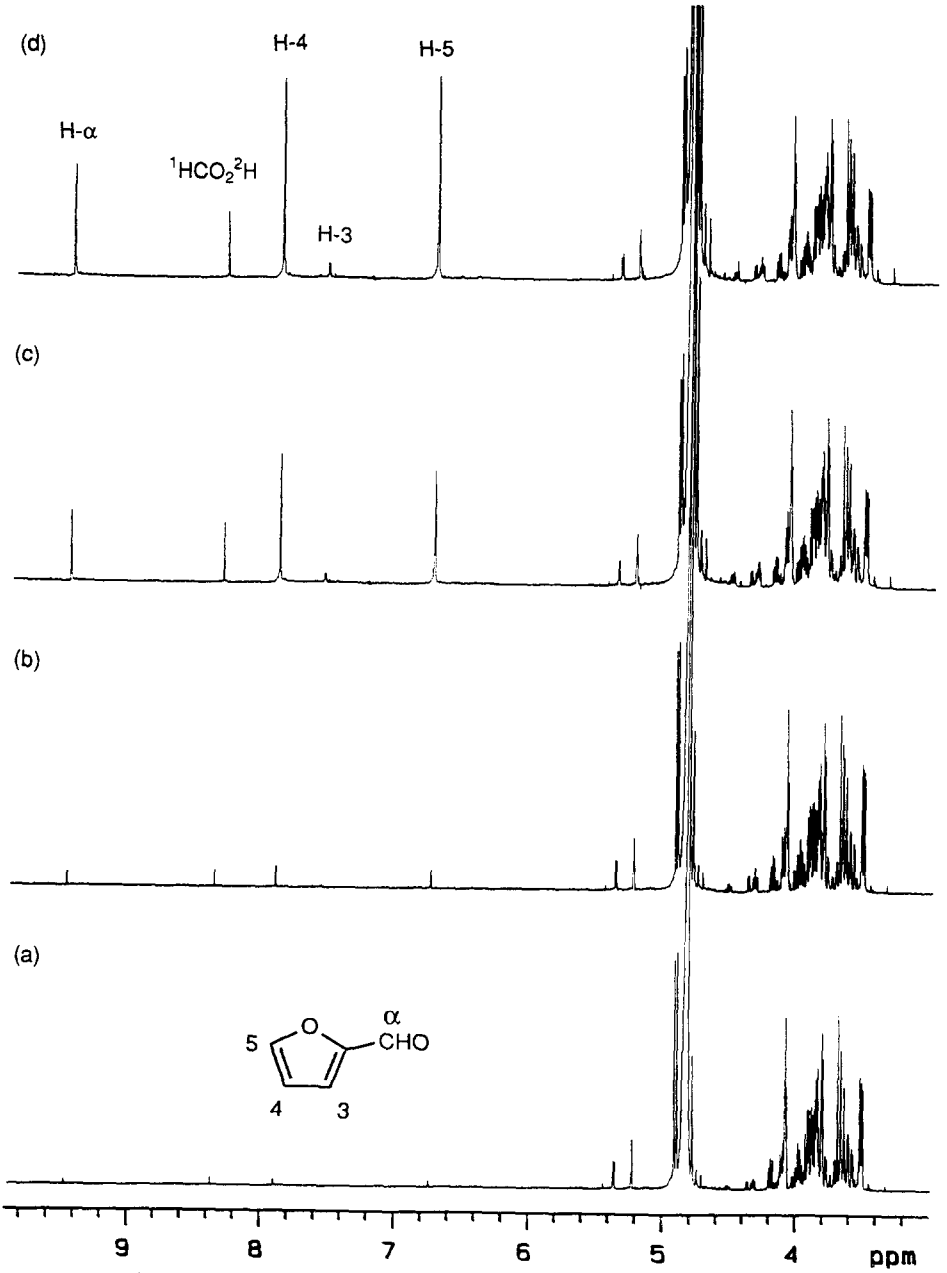


Fig. 1. 400-MHz ^{1}H NMR spectrum of D-ribose in 0.5 M deuterated sodium phosphate buffer (pD 4.5) at 25°C after heating at 96°C for (a) 2 h, (b) 8 h, (c) 48 h, and (d) 72 h.

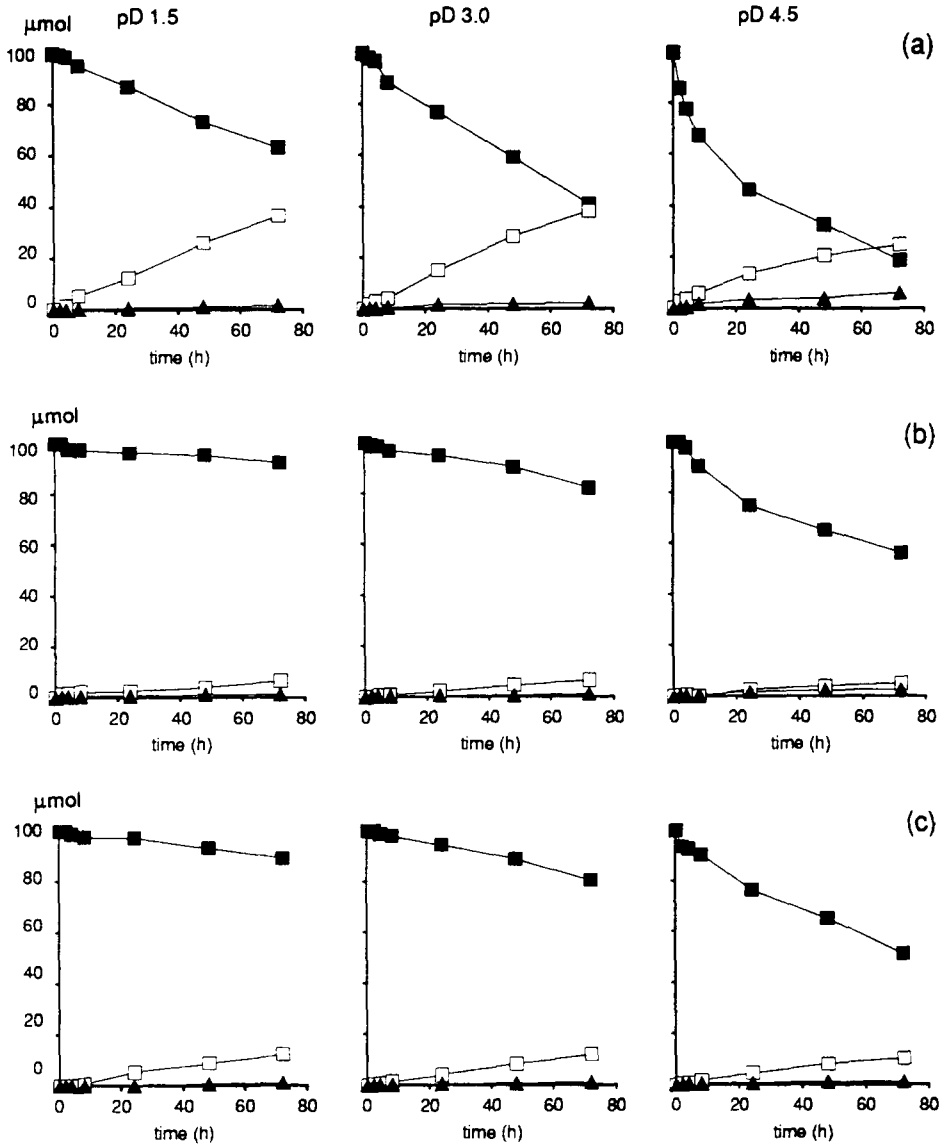


Fig. 2. Amounts of pentose (■), 2-furaldehyde (7, □), and intact formic acid ($^1HCO_2{}^2H$, ▲) after heating 100 µmol of (a) D-ribose, (b) D-arabinose, or (c) D-xylose in 0.5 M deuterated sodium phosphate buffer at 96°C.

The higher the pD, the more deuterium was incorporated and at pD 4.5 the exchange of H-3 was virtually complete.

Formic acid was formed slowly in all the degradation experiments. Notably, no

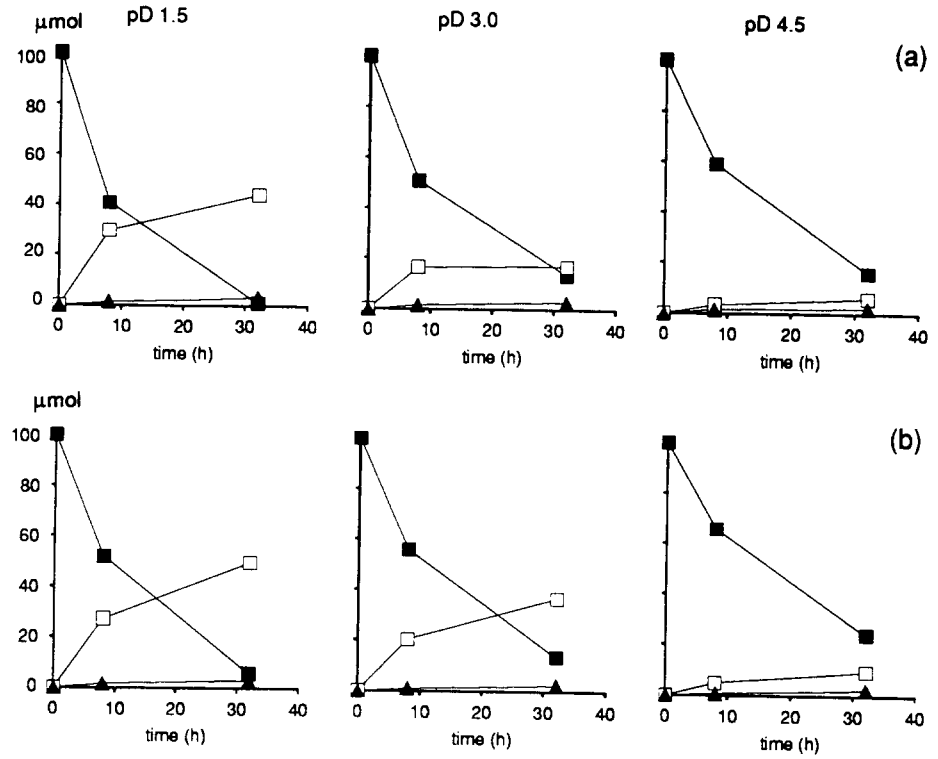


Fig. 3. Amounts of pentose (■), 2-furaldehyde (7, □), and intact formic acid ($^1HCO_2{}^2H$, ▲) after heating 100 µmol of (a) D-ribulose or (b) D-xylulose in 0.5 M deuterated sodium phosphate buffer at 96°C.

Table 1
Integrated regions in the ^1H NMR spectra of D-pentoses

D-Pentose	Proton/form	δ
Arabinose	H-2/α-*p*	3.46–3.49
Arabinose	H-1/α-*p*	4.48–4.51
Arabinose	H-1/β-*p*, H-1/α-*f*, H-1/β-*f*	5.19–5.45
Ribose [a]	H-2/β-*p*	3.49–3.54
Ribose [a]	H-2/α-*f*, H-3/α-*f*, H-3/β-*p*, H-4/α-*f*	4.05–4.15
Ribose [a]	H-1/β-*f*	5.23–5.25
Ribose [a]	H-1/α-*f*	5.36–5.38
Xylose	H-2/β-*p*, H-5b/β-*p*	3.16–3.36
Xylose	H-3/β-*p*	3.37–3.42
Xylose	H-1/β-*p*	4.53–4.56
Xylose	H-1/α-*p*	5.14–5.17
Ribulose	H-4/α-*f*, H-3/keto	4.24–4.39
Xylulose	H-3/β-*f*, H-3/α-*f*, H-4/keto	3.94–4.08
Xylulose	H-4/β-*f*	4.29–4.39

[a] See Fig. 1.

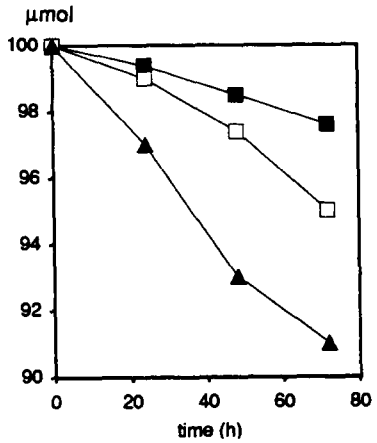


Fig. 4. Amounts of 2-furaldehyde (7) after heating 7 (100 µmol) in 0.5 M deuterated sodium phosphate buffer [pD 1.5 (▲), 3.0 (□), or 4.5 (■)] at 96°C.

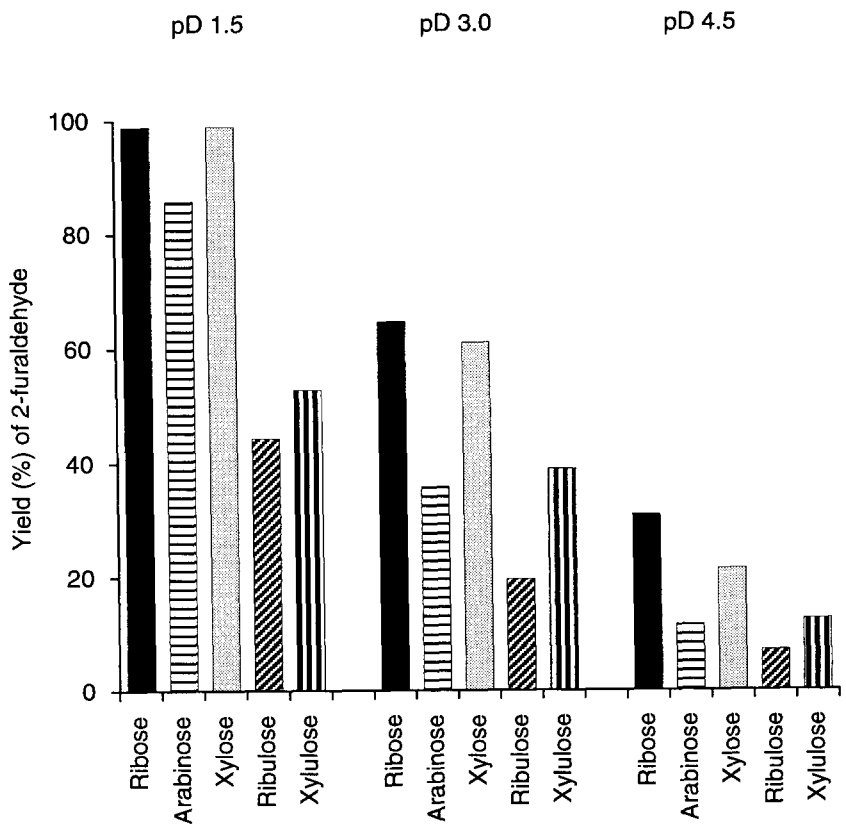


Fig. 5. Yields of 2-furaldehyde (7) calculated on consumed D-pentose at the end of the degradation experiments illustrated in Figs 2 and 3.

Table 2
^{2}H content (p_α and p_3) of H-α and H-3 in 2-furaldehyde (**7**) formed from D-pentoses in 0.5 M deuterated sodium phosphate buffer at 96°C. The data were obtained by integration of the ^{1}H NMR spectra

D-Pentose	pD	p_α (atom %)	p_3 (atom %)
Arabinose	1.5	25	100
Arabinose	3.0	39	100
Arabinose	4.5	61	100
Ribose	1.5	15	81
Ribose	3.0	42	87
Ribose	4.5	57	100
Xylose	1.5	15	98
Xylose	3.0	29	100
Xylose	4.5	50	100
Ribulose	1.5	9	79
Ribulose	3.0	39	97
Ribulose	4.5	83	100
Xylulose	1.5	8	86
Xylulose	3.0	35	96
Xylulose	4.5	74	100

formic acid was formed from **7** under any of the conditions used. For further information on the reaction mechanism, some 1- and 5-^{13}C-substituted aldopentoses were degraded for 72 h in the same way as the ordinary aldoses. The ^{1}H and the proton-decoupled ^{13}C NMR spectra of the final reaction mixtures were recorded, and the formyl signals integrated.

The isotopic composition of carbon in the formic acid ($^1HCO_2{}^2H$) with intact formyl proton was determined by integration of the ^{1}H NMR signals (a $^{12}C^1H$ singlet and a $^{13}C^1H$ doublet). The results are listed in Table 3.

Table 3
^{13}C content (p) of carbon in intact formic acid ($^1HCO_2{}^2H$) formed from ^{13}C-substituted D-pentoses in 0.5 M deuterated sodium phosphate buffer at 96°C. The data were obtained by integration of the ^{1}H NMR signals (a $^{12}C\ ^1H$ singlet and a $^{13}C\ ^1H$ doublet) from the formyl group

D-Pentose	pD	p (atom %)
[1-^{13}C] Ara	1.5	60
[5-^{13}C]Ara	1.5	50
[1-^{13}C]Ara	3.0	63
[5-^{13}C]Ara	3.0	37
[1-^{13}C]Ara	4.5	54
[5-^{13}C]Ara	4.5	52
[1-^{13}C]Rib	1.5	56
[5-^{13}C]Rib	1.5	51
[1-^{13}C]Rib	3.0	60
[5-^{13}C]Rib	3.0	48
[1-^{13}C]Rib	4.5	48
[5-^{13}C]Rib	4.5	56

From the integrals of the ^{13}C signals (a $^{13}C^{1}H$ singlet and a $^{13}C^{2}H$ triplet), the approximate deuterium content was estimated. This was 55 atom % at pD 4.5 and ≤ 35 atom % at pD 3.0, when starting from a [1-^{13}C]aldopentose. Unfortunately, the ^{13}C NMR spectra were too noisy for more precise determinations. At pD 1.5, no $^{13}C^{2}H$ signal from the formic acid was detected. This was also the case, when starting from a [5-^{13}C]aldopentose, even at pD 4.5. (No deuterium was incorporated into the formyl group when formic acid was heated for 72 h in deuterated phosphate buffer at pD 4.5.)

The highest yield of $^{1}HCO_2{}^{2}H$ (8.1 μmol from 100 μmol of consumed pentose) was obtained from ribose after 72 h at pD 4.5 (Figs 2 and 3). The part (48%) of this acid derived from C-1 of the ribose was accompanied by a somewhat larger amount of $^{2}HCO_2{}^{2}H$. Hence, the highest yield of total formic acid from 100 μmol of consumed pentose was ~ 13 μmol.

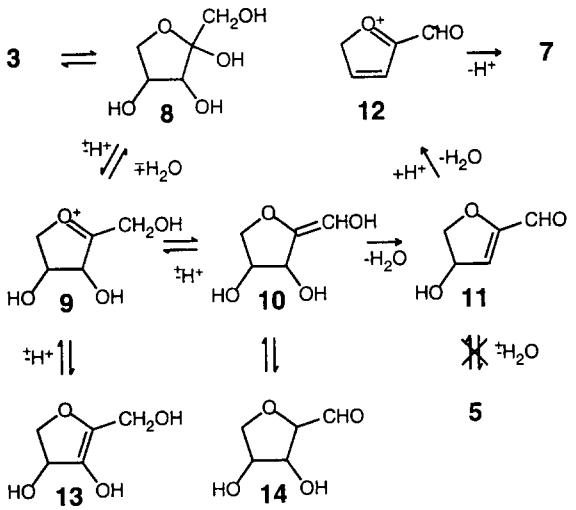
3. Discussion

For each pentose except ribulose, the integrated regions (Table 1) included signals from more than one of the isomeric forms. Since the ratios of the integrals remained constant within $\pm 5\%$ during each experiment, mutarotation was much faster than degradation and remained close to equilibrium throughout the experiment.

The results shown in Fig. 2 indicate that the stability of aldopentoses in aqueous solution is highest at a surprisingly low pH value (near or below pH 1.5). Since increasing pD accelerated their degradation but retarded the formation of 2-furaldehyde (**7**), other products [8] must have formed at a strongly increasing rate. Among these products, only formic acid was formed in amounts sufficient to permit direct quantification by ^{1}H NMR spectroscopy.

As expected from Scheme 1, the relative degradation rate of each pentose seemed to be governed by the abundance of its acyclic form at equilibrium, which in turn was inversely related to the stability of its pyranose forms. Thus, the all-*cis*-configuration of α-D-ribopyranose leads to a *syn*-1,3-diaxial interaction in both chair conformers, resulting in a lower proportion of this isomer and a higher proportion of all other isomers, including the aldehydo isomer, compared to the other aldopentoses [11]. This may explain why ribose was degraded faster than arabinose and xylose. Since no pyranose forms of the ketopentoses exist, these contain an exceptionally high proportion of keto form [12] with prominent signals in the ^{1}H NMR spectrum. Accordingly, the ketopentoses were degraded much faster than even ribose.

The data in Table 2 are also compatible with the classical route to **7** (Scheme 1). Since most or all of H-3 was exchanged in **7**, the enol **5** tautomerized faster than it was dehydrated, contrary to its behaviour in stronger acid [5]. In other words, the formation of **6** from **5** through β-elimination was slow enough to permit complete or nearly complete equilibration of **5** with **4**. Similarly, the partial exchange of H-α in **7** implies that the 1,2-enediol **2** tautomerized to a ketose (**3**) faster than it was dehydrated to **5**, although equilibrium between **2** and **3** was not attained. The gradual exchange of H-1a and H-1b in **3** confirms this interpretation. On the other hand, no hydrogen was



Scheme 2. Alternative route from pentoses to 2-furaldehyde via ketopentofuranoses.

exchanged in the aldoses (**1**). Hence, these were converted irreversibly to 1,2-enediol **2**. Since H-3 was not exchanged in **3**, any formation of 2,3-enediol from **3** was irreversible.

When "Amadori compounds" (1-amino-1-deoxyfructose derivatives) were degraded in deuterated 2 M acetic acid [15], partial exchange of H-α and H-3 in the resulting 5-(hydroxymethyl)-2-furaldehyde was observed and explained essentially as above.

Although our results are well explained by Scheme 1, they do not rule out an alternative route via the ketofuranose **8** (which is more abundant than even the keto form **3** [12]) and the cyclic intermediates **9–12** (Scheme 2). An analogous route from fructose to 5-(hydroxymethyl)-2-furaldehyde has been proposed several times [4], most recently in 1990 [16]. Scheme 2 could apply to the aldopentoses as well, if these were converted into **8** via **1–3** or into **10** via the 2,5-anhydropentoses **14** through ring contraction of the pyranoses [7].

Cyclic intermediates have been proposed partly in order to explain why no hydrogen from the solvent is found at carbon in **7** or its 5-(hydroxymethyl) derivative. However, the conclusions are based on experiments [5–7,16] performed in strong acid and/or at very high temperature. Under the milder conditions used in the present study, **10** might be dehydrated slowly enough to permit its reversible conversion into **13** via **9**, explaining the data in Table 2.

If this interpretation was correct, **10** should also have time to tautomerize to **14**. However, **14** was not observed in the ^{1}H NMR spectra. Similarly, the analogue of **14**, 2,5-anhydromannose, was not observed in the degradation of fructose and was less reactive than fructose [4]. Moreover, neither **4** nor **5** is an intermediate in Scheme 2, although **4** is no doubt involved in Maillard reactions [2] and in carbohydrate degradation under basic conditions [4]. For these reasons, we prefer to interpret our results according to Scheme 1. The route from pyranoses via **14** to **7** [7] also does not explain why ribose is degraded faster than arabinose and xylose.

The results in Table 3 indicate that the formic acid was derived exclusively from the terminal pentose carbons, C-1 being somewhat more important than C-5. The apparently exceptional data obtained at pD 4.5 are explained by the large proportion of fully deuterated formic acid ($^2HCO_2{}^2H$) at pD 4.5, for this acid was derived mainly from C-1 of the pentose [2]. Despite these results, we are still far from knowing how the formic acid is formed.

Isotope effects have been disregarded in the above discussion. The $^{13}C/^{12}C$ effects were probably not much larger than those introduced by errors in the integrals of the NMR signals. The $^2H/^1H$ effects were certainly larger but do not affect our qualitative conclusions.

4. Experimental

Materials.—All reagents and solvents were commercial samples of analytical grade. The ^{13}C-substituted pentoses were obtained from Omicron Biochemicals, Inc. Their isotopic purity was 99 atom% and that of the deuterium oxide 99.9 atom%. Orthophosphoric acid and its salts were deuterated by repeated rotary evaporation of their solutions in deuterium oxide. The buffers were prepared by adding 0.5 M D_3PO_4 or Na_2DPO_4 to 0.5 M NaD_2PO_4, until the pH meter showed the desired value (pD − 0.4) [10]. Pivalic acid (1.00 mg/mL) and sodium 3-(trimethylsilyl)propionate-d_4 (TSP, 1.00 mg/mL) were added to each buffer. This lowered the pH meter reading by 0.04.

^1H NMR spectra.—These were recorded at 400 MHz and 25°C on a Varian VXR-400 instrument and referenced to internal TSP. A 45° pulse and a 6.3-s relaxation delay were applied to allow complete relaxation. The FID data were transferred to the program NMR1 (from New Methods Research, Inc.) and then processed. The line-broadening factor was 0.3 Hz. The integrals of relevant signals were measured after performing the baseline flatting routine both automatically and manually. The integrals were calculated automatically, after the spectral lines had been identified. For each line, the integral was calculated over five line-widths. The integral phase correction was carried out manually.

^{13}C NMR spectra.—These were recorded using inverse gated decoupling, a relaxation delay of 10 s, and a pulse width of 45°. The conditions were the same as in the ^1H NMR experiments.

[2] This is exemplified by the results for ^{13}C-substituted ribose at pD 4.5 (last two lines in Table 3). For simplicity, it is assumed that the formic acid was derived from C-1 or C-5, and that all $^2HCO_2{}^2H$ was derived from C-1, as indicated by the results. Moreover, ~55 mol% of the formic acid derived from C-1 was $^2HCO_2{}^2H$. Hence, 45 mol% of it was $^1HCO_2{}^2H$. Out of 100 mol of $^1HCO_2{}^2H$ formed from [1-^{13}C]ribose, 48 mol was derived from C-1 and, hence, 52 mol from C-5. The total amount of formic acid derived from C-1 was therefore $48/0.45 = 107$ mol. Hence, $100 \times 107/(107 + 52) = 67$ mol% of the formic acid was derived from C-1 and 33 mol% from C-5. Similarly, out of 100 mol of $^1HCO_2{}^2H$ formed from [5-^{13}C]ribose, 56 mol was derived from C-5 and, hence, 44 mol from C-1. The total amount of formic acid derived from C-1 was therefore $44/0.45 = 98$ mol. Hence, $100 \times 98/(98 + 56) = 64$ mol% of the formic acid was derived from C-1 and 36 mol% from C-5. Without more accurate results on the deuterium content of the formic acid, we made no similar calculations from the other data in Table 3.

Degradation experiments.—In each experiment, a pentose (15.0 mg, 100 μmol) or **7** (9.6 mg, 100 μmol) was dissolved in the appropriate buffer (0.60 mL) and transferred to a 5-mm NMR tube. The tube was sealed, the ^{1}H NMR spectrum recorded, and the tube heated in a water bath at 96°C. At certain times (t) the tube was cooled quickly. After 1–2 h at 25°C, the ^{1}H NMR spectrum was recorded again, and the heating continued immediately. When an aldose was degraded, t was 2, 4, 8, 24, 48, and 72 h; when a ketose was degraded, t was 8 and 32 h; and when **7** was degraded, t was 24, 48, and 72 h. The heating times t are indicated in Figs 1–4 and do not include the waiting times at 25°C. The data in Table 2 were obtained from the final spectra (after 32 or 72 h) and have been corrected for the ^{1}H content (0.7 atom%) of the medium (mainly due to pentose-O^1H). When the [^{13}C]aldopentoses were degraded, ^{1}H and ^{13}C NMR spectra were recorded only after 72 h. The data in Table 3 have been corrected for the ^{12}C content (1 atom%) of the ^{13}C-substituted carbons.

Acknowledgements

We thank Mr Rolf Andersson for expert advice on recording the NMR spectra. The present work was supported by grants from the Swedish Natural Science Research Council and the Swedish Farmers' Foundation for Agricultural Research.

References

[1] O. Theander and D.A. Nelson, *Adv. Carbohydr. Chem. Biochem.*, 46 (1988) 273–326.
[2] F. Ledl and E. Schleicher, *Angew. Chem. Int. Ed. Engl.*, 29 (1990) 565–594.
[3] F.A. Riera, R. Alvarez, and J. Coca, *J. Chem. Technol. Biotechnol.*, 50 (1991) 149–155.
[4] M.S. Feather and J.F. Harris, *Adv. Carbohydr. Chem. Biochem.*, 28 (1973) 161–224.
[5] M.S. Feather, D.W. Harris, and S.B. Nichols, *J. Org. Chem.*, 37 (1972) 1606–1608.
[6] D.W. Harris and M.S. Feather, *Carbohydr. Res.*, 30 (1973) 359–365; *J. Org. Chem.*, 39 (1974) 724–725.
[7] M.J. Antal, Jr, T. Leesomboon, W.S. Mok, and G.N. Richards, *Carbohydr. Res.*, 217 (1991) 71–85.
[8] T. Popoff and O. Theander, *Carbohydr. Res.*, 22 (1972) 135–149; D.A. Nelson, R.T. Hallen, and O. Theander, *ACS Symp. Ser.*, 376 (1988) 113–118.
[9] I. Forsskåhl, T. Popoff, and O. Theander, *Carbohydr. Res.*, 48 (1976) 13–21.
[10] P.E. Pfeffer, K.M. Valentine, and F.W. Parrish, *J. Am. Chem. Soc.*, 101 (1979) 1265–1274.
[11] S.J. Angyal, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 15–68.
[12] J. Wu, A.S. Serianni, and T. Vuorinen, *Carbohydr. Res.*, 206 (1990) 1–12.
[13] A.J. Benesi, C.J. Falzone, S. Banerjee, and G.K. Farber, *Carbohydr. Res.*, 258 (1994) 27–33; T. Vuorinen and A.S. Serianni, *ibid.*, 209 (1990) 13–31.
[14] S.J. Angyal and V.A. Pickles, *Aust. J. Chem.*, 25 (1972) 1695–1710.
[15] M.S. Feather, *Prog. Food Nutr. Sci.*, 5 (1981) 37–45.
[16] M.J. Antal, Jr, W.S.L. Mok, and G.N. Richards, *Carbohydr. Res.*, 199 (1990) 91–109.