ELSEVIER

www.elsevier.nl/locate/carres

Carbohydrate Research 329 (2000) 359–365

CARBOHYDRATE RESEARCH

A kinetic estimate of the free aldehyde content of aldoses

Jason P. Dworkin [1], Stanley L. Miller *

Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0506, USA

Received 19 July 1999; accepted 20 June 2000

Abstract

The relative free aldehyde content of eight hexoses and four pentoses has been estimated within about 10% from the rate constants for their reaction with urazole (1,2,4-triazole-3,5-dione). These values of the percent free aldehyde are in agreement with those estimated from CD measurements, but are more accurate. The relative free aldehyde contents for the aldoses were then correlated to various literature NMR measurements to obtain the absolute values. This procedure was also done for three deoxyaldoses, which react much more rapidly than can be accounted for by the free aldehyde content. This difference in reactivity between aldoses and deoxyaldoses is due to the inductive effect of the H versus the OH on C-2′. This may help explain why deoxyribonucleosides hydrolyze much more rapidly than ribonucleosides. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Free aldehyde; Aldoses; Urazole; Depurination

1. Introduction

The free aldehyde content (fraction of free aldehydo form in aqueous solution of an aldose in mutarotational equilibrium) of aldoses is important in understanding aldose reactivity [1,2] and stability [3]. It may also be related to the choice of glucose as the major hexose in biology, since it has the least free aldehyde and is thus the least reactive hexose with proteins [4]. This choice has been attributed [5] to the all-equatorial hydroxyls of glucose, but this seems unlikely as a dominant factor since the difference in free energy of formation of the various hexoses is small (~ 1.2 kcal/mol) [6]. However, the very low free aldehyde content of aldoses is very difficult to measure.

The first accepted measurement for the free aldehyde content of glucose is 0.0026%, which was determined via polarography at 25 °C [7]. However, there is only one survey of relative aldehyde content for the straight chain aldopentoses and aldohexoses. This survey (at 20 °C) used the circular dichroism (CD) of the 290 nm n to π* transition in the carbonyl functional group, calibrated from the optical rotatory dispersion (ORD) of various ketones and aldehydes in methanol, heptane, and chloroform [8], and was confirmed by the agreement of the glucose value with the polaragraphic value. Angyal [9] stated that these values are correct within an order of magnitude and are primarily reliable only on a relative basis. Subsequent measurements of ^{13}C-enriched glucose via ^{13}C NMR spec-

* Corresponding author. Tel.: + 1-858-5343365; fax: + 1-858-5346128.

E-mail address: smiller@ucsd.edu (S.L. Miller).

[1] Present address: NASA Ames Research Center, M/C 245-6, Moffett Field, CA 94035-1000, USA.

0008-6215/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0008-6215(00)00204-4

troscopy can be extrapolated to 20 °C to give a substantially different value [10]. NMR values for idose, ribose, and talose were also considerably different [1,11,12] from those obtained from CD.

In the course of investigating the reactivity of ribose and other aldoses with urazole (1,2,4-triazole-3,5-dione) [13,14], it became apparent that the rates are approximately proportional to the content of free aldehyde. We have measured these relative values based on kinetics and calibrated the absolute free aldehyde values with several NMR measurements. This is useful as a systematic measurement of the free aldehyde, particularly since the free aldehyde content of most aldoses has not been measured by NMR spectroscopy. We have also measured 2-deoxyribose (i.e., 2-deoxy-D-*erythro*-pentose) and two of the 2-deoxyhexoses. These results allow us to constrain a model that can explain the rapid acid-catalyzed hydrolysis of deoxynucleosides relative to ribonucleosides. The rapid hydrolysis of purine deoxyribosides is a major source of damage to DNA for which various repair systems are needed [15].

2. Experimental

Materials.—1-Methylurazole was synthesized by the method of Bausch et al. [16]. Urazole and 4-methylurazole were from Aldrich Chemical Co. Sigma Chemical Co. supplied the glycolaldehyde, paraldehyde, dihydroxyacetone, and all the sugars except for gulose (Biospherics) and glucose (Fisher). All deuterated solvents were from Cambridge Isotope Laboratories. The acetaldehyde (Mallinckrodt) was purified by vacuum distillation from $CHCl_3$ (l)/$CHCl_3$ (s) to N_2 (l) to remove any paraldehyde or other contaminants immediately before use. The formaldehyde was distilled from paraformaldehyde in D_2O and H_2SO_4, and the concentration was determined by a bisulfite titration [17].

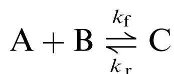
Instrumentation.—NMR analysis was done on a GE QE-300 NMR with a 5-mm broad band probe. Reaction temperatures were held constant (± 0.2 °C as measured by a quartz thermometer) in thermostated baths or in the variable-temperature unit on the NMR instrument. HPLC analysis was with two Beckman 110B pumps controlled by Beckman System Gold and analyzed by a Kratos Spectroflow 757 UV detector set to 232 nm. Riboside separation was achieved with an Alltech HEMA-IEC BIO 1000 Q 10U (10×250 mm) strong anion-exchange semipreparative column. The solvent system was 2.0 mM NaOH at a flow rate of 4 mL/min.

Reaction of urazole with acetaldehyde.—A 0.1 M soln of urazole (0.5 mL) was frozen with N_2(l) and degassed in an NMR tube. Enough acetaldehyde to make a 0.1 M soln was then distilled into the NMR tube. The tube was sealed and stored in N_2(l) until the NMR spectrum could be taken (generally within the hour). The soln was melted immediately before being placed in the NMR instrument.

Reaction of urazoles with ribose and other aldoses.—Urazole and D-ribose (0.2 M in each) were dissolved in D_2O and sealed in an NMR tube. The mixture was heated at various temperatures (25, 40, 60, 80, 100, and 120 °C) and times (up to 2 years). The pD was measured to be approximately 4.5, where the rate constants are independent of pH [13,14]. The progress of the reaction was monitored by ^1H NMR spectroscopy. The products of the reaction with ribose were separated via preparative scale HPLC, and the individual peaks were characterized via ^1H and ^{13}C NMR spectroscopy. Glycolaldehyde, glyceraldehyde, 2-deoxyribose, 2-deoxygalactose, 2-deoxyglucose and the other 11 straight-chain aldopentoses and aldohexoses were treated similarly. The ^1H NMR spectra of the reactions with 2-deoxyribose and 2-deoxygalactose were taken at 80 and 60 °C, respectively. This was necessary to shift the β-P peaks of the aldose away from the HOD peak. Once the sealed NMR tube was analyzed in this fashion, it was not used again. Thus, unlike the profiles generated by the reaction of the other aldoses, each point represents a separate sample.

The rate constants were determined by following the reaction with ^1H NMR spectroscopy and taking measurements of the anomeric protons at various times. These

measurements were then fitted with the second-order reversible rate equation [18] (where $[A]_0 = [B]_0$ and $[C]_0 = 0$) to determine the rate constants and their standard deviations:



$$A + B \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C$$

$$K = \frac{[C]_\infty}{[A]_\infty [B]_\infty} = \frac{k_f}{k_r}$$

$$[A]_\infty = \frac{1}{2K}(\sqrt{1 + 4K[A]_0} - 1)$$

$$k_f = \frac{K}{t(2K[A]_\infty + 1)}$$

$$\ln \frac{([A] + [A]_\infty + K^{-1})([A]_0 - [A]_\infty)}{([A]_0 + [A]_\infty + K^{-1})([A] - [A]_\infty)}$$

Free aldehyde content of 2-deoxyribose.— The free aldehyde peak (broad peak at 10.4 ppm) of a 1.0 M soln of deoxyribose was observed by ^{1}H NMR at 100, 80, 60, and 40 °C and integrated against the anomeric protons. The soln was allowed to equilibrate at the appropriate temperature before acquisition.

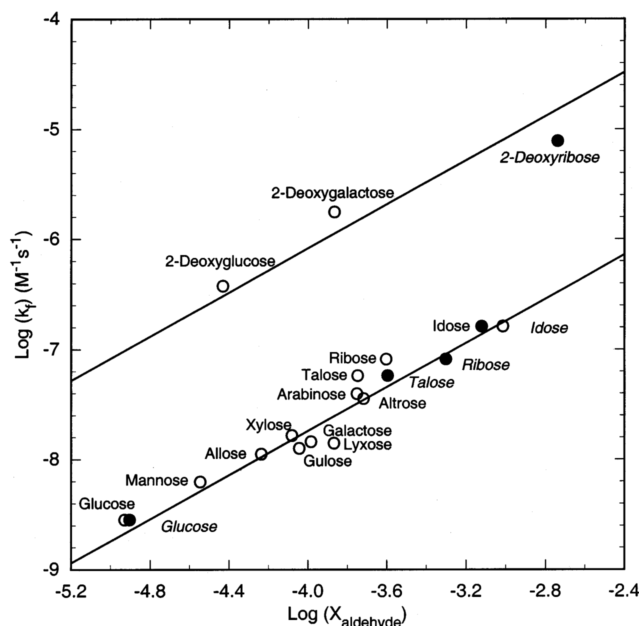


Fig. 1. Second-order rate constant for the formation of all four aldosides versus mole fraction free aldehyde ($X_{aldehyde}$). The closed circles with labels in italics have had their $X_{aldehyde}$ determined by NMR; the open circles have had their $X_{aldehyde}$ calculated as described in the text.

3. Results

The reaction of urazole with aldoses gave reversible first-order kinetics [14]. The equilibrium constants $K_{eq} = [UR]/[U][R]$ varied between 0.5 and 3.1 [14]. Reaction rates with various aldoses were monitored over time at 25 °C except for mannose and glucose, which were obtained by Arrhenius extrapolations from data at 120, 100, 80, and 60 °C. Some of the aldoses were also measured at 100 °C in addition. The reaction was determined to be at the N_1 of the urazole [13], and the rate of ring closure of the aldoside is unlikely to be the slow step, since the reaction of model compounds (acetaldehyde and glycolaldehyde) with urazole reached an equilibrium within 2 min at 20 °C. The K_{eq} for acetaldehyde for the single addition (N_1) is 18.3, and the K_{eq} is 1.3 for the double addition (N_1 and N_2). The reaction of acetaldehyde with 4-methylurazole has the same K_{eq} and k as with urazole, and the same double addition products are observed. The reaction of acetaldehyde with 1-methylurazole produces the single addition with the same equilibrium and rate constants as the single addition of urazole. This shows the similar steric hindrance of the methyl and hydroxyethyl groups. The reaction with formaldehyde gave single (5.0 ppm) and double (5.1 ppm) adducts in a 2:1 ratio with a K_{eq} for the single adduct of about 600.

The equilibrium constant for the reaction with glyceraldehyde was determined from ratios of the aldehyde peak (9.65 ppm) to the peaks in the anomeric region of the spectra with and without urazole. The K_{eq} was estimated to be 7 ± 1 in terms of total glyceraldehyde. The reaction with glycolaldehyde was more easily quantitated and the K_{eq} was determined to be 5.1 ± 0.2.

Urazole does not react detectably ($< 5\%$) under the conditions investigated with the ketoses tested (fructose, sorbose and tagatose), nor does it react with acetone or dihydroxyacetone ($< 1\%$). Thus K_{eq} is < 0.1 for both the ketoses and the ketones.

Fig. 1 shows the rate constants plotted against the mole fraction of the free aldehyde based on the CD data [8]. Three adjustments have been made to these data:

Table 1
Comparison of CD and kinetic data for aldose free aldehyde content

Aldose	k (M^{-1} s^{-1}) $\times 10^9$ at 25 °C [a]	% Aldehyde $\times 10^3$		
		From CD at 20 °C [b]	Adjusted [c]	Urazole data [d]
Idose	160 ± 20	182	96.9	78.6
Ribose	81.6 ± 0.7	46.9	25.0	39.6
Talose	58 ± 2	33.6	19.2	28.0
Arabinose	39.6 ± 0.5	33.3	17.9	19.2
Altrose	36 ± 1	36.0	17.7	17.3
Xylose	16.5 ± 0.7	15.6	13.6	7.97
Gulose	14.4 ± 0.2	19.5	10.4	7.00
Lyxose	14.0 ± 0.3	25.5	9.05	6.79
Galactose	12.6 ± 0.6	17.0	8.30	6.10
Allose	11.2 ± 0.4	10.9	5.80	5.41
Mannose	6.2 ± 1	5.35	2.85	3.04
Glucose	2.8 ± 0.4	2.22	1.18	1.38

[a] Second-order rate constant, e.g., the reaction of glucose with urazole is 2.8×10^{-9} M^{-1} s^{-1}.
[b] Haywood and Angyal CD data using the three significant figures of the measurement. The value for gulose is taken from the corresponding heptose (D-*glycero*-D-*gluo*-heptose).
[c] CD data adjusted to 25 °C and to the NMR data as described in the text.
[d] % aldehyde from the urazole kinetic data normalized to the ^{13}C NMR free aldehyde data. These figures can be derived from $\log X_{aldehyde} = a + \log k$ where k is the rate constant shown in the first column and a is the normalization to the NMR data of 3.686.

1. The three significant figures listed [8] for the CD measurements were used to estimate the percentage aldehyde instead of the rounded figures recommended [8] (glucose = 2.22×10^{-3}% rather than = 2×10^{-3}%).
2. The 20 °C data[2] (glucose = 2.22×10^{-3}%) have been adjusted to 25 °C (glucose = 2.96×10^{-3}%) by multiplying by 1.33, based on $\Delta H = 9.98$ kcal/mol from the NMR measurement of glucose [10].
3. These CD free aldehyde values were then adjusted so that they are in agreement with a least-squares fit of the ^{13}C NMR (adjusted to 25 °C) values for ribose (0.050%) [12], glucose (1.25×10^{-3}%) [10], idose (0.0757%) [11], and talose (0.0254%) [1]. This adjustment gives glucose = 1.18×10^{-3}%.

The results of these adjustments for all the aldoses investigated are shown in Table 1, column 4.

[2] These data are reported for 20 °C and rounded off to one significant figure. We have used all three figures of the CD data.

Aldoses.—Rather than fitting the CD values [8] to the various NMR data, the rate constant for the reaction of aldoses with urazole can be fitted to the NMR data. The use of the kinetic data to estimate the percent free aldehyde is valid with the assumption that the reaction rate is proportional to the mole fraction of free aldehyde. These values are also shown in Table 1 as well at the rate data with the error given as 1σ.

The error was based on the standard deviations of the rate constant calculated for each individual time point for a given reaction (typically 24 data points taken over 2 years of reaction where $t_{1/2} = 85$ days for idose and about 6.7 years for glucose). This gives us a better measurement of the error than that of the CD measurements [8]. While we have found that the relative errors of the CD technique were generally small, the kinetic approach gives a superior measurement of precision with at least as good accuracy (if calibrated via NMR spectroscopy) as other techniques.

2-Deoxyaldoses.—It has been reported that urazole and deoxyribose react quickly [19,20]. We have measured the rate constants for de-

oxyribose and two deoxyhexoses with urazole at 25 °C. The deoxyaldose rate constants are also shown in Fig. 1. These data, although less extensive, clearly show that the deoxyaldoses sit on a separate line. The NMR data for the free aldehyde content of 2-deoxyribose (Fig. 2) was used in Fig. 1. The equation of the line in Fig. 2 is

$$\log(X_{\text{aldehyde}}) = -\frac{1360}{T} + 1.82$$

This gives the free aldehyde content at 25 °C as 0.18% and a ΔH of 6.2 kcal/mol, which is to be compared with 9.98 for glucose [10]. This is also in agreement with the 0.3% aldehyde content of deoxyribose at 43.3 °C reported by Bauer et al. [21]. Since the CD data [8] only contain 2-deoxyglucose and 2-deoxygalactose and not 2-deoxyribose, we have not attempted to adjust the CD deoxyaldose data to the line generated by the NMR measurement of the deoxyribose.

4. Discussion

The kinetic procedure to obtain the relative percent free aldehyde can be justified on the following A_2 reaction mechanism. This has also been proposed for the depyrimidation of 2′-deoxyuridine and thymidine [22] (Scheme 1).

The K_{op} and K_{add} are stated as in rapid equilibrium and the k_3 is the slow step[3]. If we assume K_{add} and k_3 are the same for all aldoses, then the ratio of the rates is given by the ratios of K_{op} (the amount of free aldehyde). The inductive effect of N would be nearly the same for all aldoses since they differ only in the relative configuration of the hydroxyls.

While there are only three points, the deoxyaldoses are on a different line from the aldoses, which is faster than the aldose line by a factor of 50. This difference should be in k_3 and/or in K_{add}. Our measured value of K_{add} for aqueous acetaldehyde (a model for open chain 2-deoxyaldoses) is 38 in terms of the measured free aldehyde. The K_{add} for aqueous glycolaldehyde (a model for open chain aldoses) is 91 in terms of the measured free aldehyde.

We take

$$\frac{K_{eq}^{\text{glycoaldehyde}}}{K_{eq}^{\text{acetaldehyde}}} = \frac{K_2^{\text{aldose}}}{K_2^{\text{deoxyaldose}}}$$

Since the inductive effect of the aldose component in compound A is the same for all the deoxyaldoses, the line in Fig. 1 for the deoxyaldoses should lie below the aldoses by a factor of 2.4 while it lies above the line by a factor of 50. Therefore the difference in the rate of reaction of urazole with deoxyaldoses and with aldoses is a factor of 120. We attribute this to differences in k_3.

The rate of reaction of aldoses with urazole parallels their reaction with hemoglobin [4]. The reasonable correlation supports the conclusion that like the reaction of aldoses with protein, the rate of reaction of urazole with aldoses is based on the free aldehyde content of the aldose. Likewise, the rate of decomposition of aldoses roughly follows the free aldehyde content [3], although there is considerably more scatter.

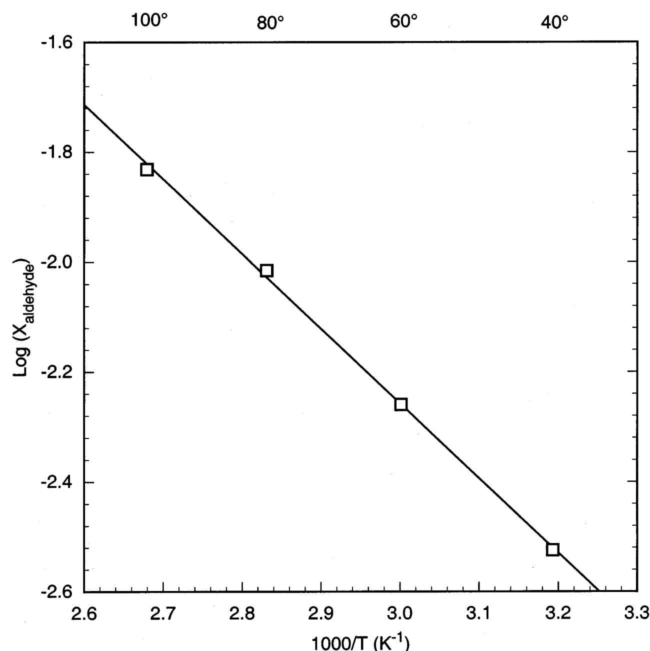
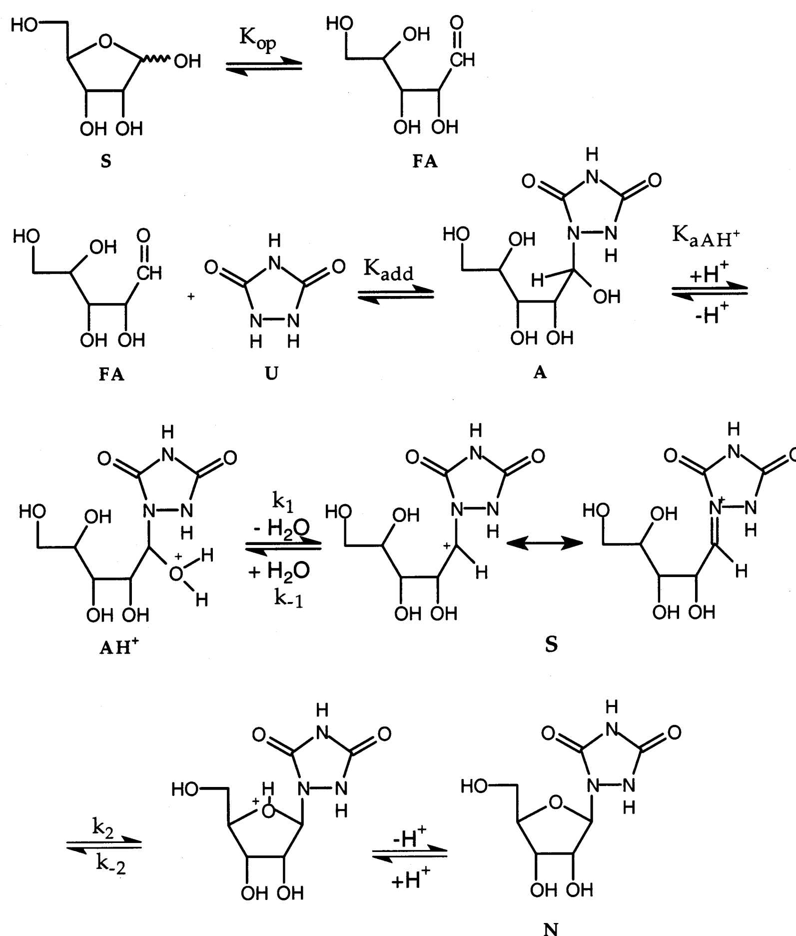
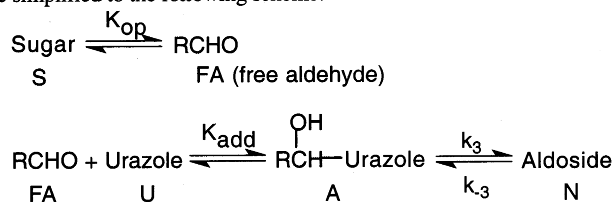


Fig. 2. Mole fraction of free aldehyde in 2-deoxyribose versus temperature based on determinations of 1.0 M deoxyribose in D_2O. The temperature was controlled with the variable temperature unit in the GE QE-300 NMR.

[3] The ratio of the rate of reaction with arabinose, lyxose, allose, galactose, glucose, mannose, and talose at 100 and 25 °C with ribose were compared and the $\Delta\Delta H$ calculated for the aldoses sits within experimental error.



This can be simplified to the following scheme:



Scheme 1.

It is possible to use the difference in k_3 for the reaction of urazole with aldoses and 2-deoxyaldoses to speculate [14] on the rates of hydrolysis of deoxynucleosides and nucleosides, which is the reverse reaction. The similarity of the ratio of k_3 values for deoxyaldoses/aldoses to the value for deoxynucleosides/nucleosides could be useful in understanding the instability of DNA under acidic conditions [23–26]. A major pathway of DNA decomposition is depurination and depyrimidination [16]. This would have been even more important before sophisticated repair enzymes had been developed, soon after the origin of life. It is possible that if the genetic material that immediately preceded the RNA of the RNA world was aldose based, it would have used a more electron-withdrawing substituent on C-2′ of pentoses or hexoses [27] to reduce this problem.

Acknowledgements

We thank Professor Charles Perrin for helpful comments and Michael Robertson and Matthew Levy for helpful comments and technical assistance and the NASA Specialized Center of Research and Training in exobiology for a fellowship (J.P.D.) and grant support (S.L.M.).

References

[1] J.R. Snyder, E.R. Johnston, A.S. Serianni, *J. Am. Chem. Soc.*, 111 (1989) 2681–2687.

[2] J.A. Wilde, P.H. Bolton, A. Mazumder, M. Manoharan, J.A. Gerlt, *J. Am. Chem. Soc.*, 111 (1989) 1894–1897.

[3] R. Larralde, M.P. Robertson, S.L. Miller, *Proc. Natl. Acad. Sci. USA*, 92 (1995) 8158–8160.

[4] H.F. Bunn, P.J. Higgins, *Science*, 213 (1981) 222–224.

[5] J. Hirabayashi, *Q. Rev. Biol.*, 71 (1996) 365–380.

[6] R.N. Goldberg, Y. Tewari, *J. Phys. Chem. Ref. Data*, 18 (1989) 809–880.

[7] J.M. Los, L.B. Simpson, K. Wiesner, *J. Am. Chem. Soc.*, 78 (1956) 1564–1569.

[8] L.D. Hayward, S.J. Angyal, *Carbohydr. Res.*, 53 (1977) 13–20.

[9] S.J. Angyal, *Adv. Carbohydr. Chem. Biochem.*, 42 (1984) 15–68.

[10] S.R. Maple, A. Allerhand, *J. Am. Chem. Soc.*, 109 (1987) 3168–3169.

[11] J.R. Snyder, A.S. Serianni, *J. Org. Chem.*, 51 (1986) 2694–2702.

[12] M.J. King-Morris, A.S. Serianni, *J. Am. Chem. Soc.*, 109 (1987) 3501–3508.

[13] V.M. Kolb, J.P. Dworkin, S.L. Miller, *J. Mol. Evol.*, 38 (1994) 549–557.

[14] J.P. Dworkin, *Ph.D. Dissertation*, University of California, San Diego, 1997.

[15] E.C. Friedberg, G.C. Walker, W. Siede, *DNA Repair and Mutagenesis*, ASM Press, Washington, DC, 1995, pp. 13–14.

[16] M.J. Bausch, B. David, P. Dobrowolski, C. Guadalupe-Fasano, R. Gostowski, D. Selmarten, V. Prasad, A. Vaughn, L.-H. Wang, *J. Org. Chem.*, 56 (1991) 5643–5651.

[17] J.F. Walker, *Formaldehyde*, Reinhold, New York, 1964, pp. 486–488.

[18] J.H. Espenson, *Chemical Kinetics and Reaction Mechanisms*, McGraw-Hill, New York, 1981, pp. 45–48.

[19] P.D. Robinson, C.Y. Meyers, V.M. Kolb, P.A. Colloton, *Acta Crystallogr., Sect. C*, 52 (1996) 1215–1218.

[20] V.M. Kolb, C.E.C.A. Hop, *Abstracts of Papers*, 210th National Meeting, American Chemical Society, Chicago, IL, August 1995, CARB-3.

[21] H. Bauer, A. Brinkmeier, J. Buddrus, M. Jablonowski, *Justus Liebigs Ann. Chem.*, (1986) 1804–1807.

[22] J. Cadet, R. Teoule, *J. Am. Chem. Soc.*, 96 (1974) 6517–6519.

[23] L. Hevesi, E. Wolfson-Davidson, J.B. Nagy, O.B. Nagy, A. Bruylants, *J. Am. Chem. Soc.*, 94 (1972) 4715–4720.

[24] E.R. Garrett, J.K. Seydel, A.J Sharpen, *J. Org. Chem.*, 31 (1966) 2219–2227.

[25] E.R. Garrett, P.J. Mehta, *J. Am. Chem. Soc.*, 94 (1972) 8532–8541.

[26] T. Fujii, T. Saito, K. Iguchi, *Chem. Pharm. Bull.*, 42 (1994) 495–499.

[27] A. Eschenmoser, *Orig. Life Evol. Biosph.*, 27 (1997) 535–553.