Published in final edited form as:

J Phys Org Chem. 2015 March ; 28(3): 173–180. doi:10.1002/poc.3318.

Hydrolytic Fitness of *N*-glycosyl Bonds: Comparing the Deglycosylation Kinetics of Modified, Alternative and Native Nucleosides

Andro C Rios, Hiu T Yu, and Yitzhak Tor¹

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0358

Abstract

Nature's selection of the contemporary nucleobases in RNA and DNA continues to intrigue the origin of life community. While the prebiotic synthesis of the *N*-glycosyl bond has historically been a central area of investigation, variations in hydrolytic stabilities among the *N*-glycosyl bonds may have presented an additional selection pressure that contributed to nucleobase and nucleoside evolution. To experimentally probe this hypothesis, a systematic kinetic analysis of the hydrolytic deglycosylation reactions of modified, alternative and native nucleosides was undertaken. Rate constants were measured as a function of temperature (at pH 1) to produce Arrhenius and Eyring plots for extrapolation to 37°C and determination of thermodynamic activation parameters. Rate enhancements based on the differences in reaction rates of deoxyribo- and ribo-glycosidic bonds were found to vary under the same conditions. Rate constants of deoxynucleosides were also measured across the pH range of 1 – 3 (at 50°C), which highlighted how simple changes to the heterocycle alone can lead to significant variation in deglycosylation rates. The contemporary nucleosides exhibited the slowest deglycosylation rates in comparison to the non-native/alternative nucleosides, which we suggest as experimental support for nature's selection of the fittest *N*-glycosyl bonds.

Keywords

Chemical evolution; nucleoside; hydrolysis; Genetic alphabet; RNA world; prebiotic chemistry

INTRODUCTION

Questions surrounding the natural selection of the canonical nucleobases are fundamental to understanding the chemical evolution of nucleic acids as well as the origin of the genetic alphabet ^{[1],[2]} Many investigations have centered on the synthetic prebiotic chemistry of the native bases and related heterocycles including their hydrolytic and photochemical

¹Yitzhak Tor, Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, 858-534-6401, ytor@ucsd.edu.

SUPPORTING INFORMATION

Additional supporting information, including spectroscopic monitoring of deglycosylation reactions, Eyring plots and tables, may be found in the online version of this article at the publisher's website.

stabilities.^[3] Understandably, a much greater focus has been placed on the particular base pairing properties and subsequent duplex stabilities using canonical and non-canonical bases along with alternative backbones.^[4–6]

What about the hydrolytic stability of the *N*-glycosidic bonds? These vital chemical linkages have also been the subject of related investigations, but largely in the context of their prebiotic chemical formation.^[7–12] Surprisingly, little attention has been given to the specific role of *N*-glycosyl stability as a possible driver in the natural selection of the native bases.^[1,9,11,13] Nucleoside glycosidic bond hydrolysis data have indeed been gathered over the years,^[14–19] but mostly in the context of nucleic acids damage, mutagenesis and repair.^[16,20–21] Some of those previous mechanistic investigations have exposed differences in *N*-glycosyl stabilities among nucleosides of closely related heterocycles. These observations led us to suggest a scenario, where the DNA alphabet letters may have been selected, in part, as a result of their higher *N*-glycosyl stabilities compared to alternative letters.^[22] To experimentally pursue this premise we systematically study here the hydrolytic stability of *N*-glycosyl bonds in selected nucleosides (Figure 1, Table 1). This subgroup represents prebiotically relevant heterocycles related to RNA and DNA base alternatives, and ultimately reflects on the question surrounding the origin of the extant genetic alphabet.^[23–24]

Deglycosylation kinetics were determined for ribo- and deoxyribo-nucleosides to address the following fundamental questions: a) do the relative stabilities of the *N*-glycosyl bonds follow a predicted outcome under the conditions examined? b) does the stability difference between a ribo-glycosidic and a deoxyribo-glycosidic bond remain the same or does it vary depending on the nature of the attached nucleobase?, and c) how do the native nucleosides, with a particular focus on the deoxynucleosides, compare in their *N*-glycosyl stability to other related modified or alternative nucleosides? Experimentally, the temperature dependence of deglycosylation rates was measured under standardized acidic conditions (0.1M HCl) for all nucleosides in order to obtain rate extrapolations, enhancements and thermodynamic activation parameters at 37°C from the corresponding Arrhenius and Eyring plots. Additionally, comparative hydrolysis kinetics at 50°C between pH 1 and pH 3 was determined for the deoxynucleosides. Low pH was chosen for this study mainly to facilitate the comparison of ribonucleoside and deoxyribonucleoside kinetics under the same experimental conditions and for comparison to previously published studies.

We observed that the heterocycles' protonation sites as determined from their pK_a values, in addition to their electronic nature, had the greatest impact on relative deglycosylation rates. We also observed that, in comparing the ribo- and deoxyribonucleoside series, deglycosylation rate enhancements were dependent on the identity of the heterocycle. Lastly we found that non-native/alternative nucleosides tend to exhibit faster deglycosylation rates in comparison to the native nucleosides suggesting their lower evolutionary fitness level.

EXPERIMENTAL SECTION

Nucleosides and nucleobases were purchased from Sigma-Aldrich, Acros, or MP Biochemicals and used without further purification. 2'-Deoxyisoguanosine, was purchased

from Berry & Associates (Dexter, Michigan). 2'-Deoxyxanthosine was purchased from Carbosynth LLC (San Diego, CA). Isoguanosine, isoguanine, 2,6-diaminopurine-2'-deoxyriboside, and zebularine were purchased from Toronto Research Chemicals (North York, Ontario). 2'-Deoxyzebularine was purchased from Trilink Biotechnologies (San Diego, California). For experiments conducted under acidic conditions (pH 1 – 3) and 10 – 85 °C temperature range, a Shimadzu UV-2450, UV-Vis Spectrophotometer was used to obtain kinetics in real-time or after an alkaline quench (as described below) by monitoring at a single wavelength from the observed spectra. Rate constants were determined by fitting the resulting pseudo-first order decay curves using Origin Pro 9.0 graphing software (OriginLab Corporation, North Hampton, MA).

2'-Deoxyribonucleosides

Reaction mixtures for deglycosylation kinetics of deoxynucleosides at pH 1 (0.1M HCl) were prepared in 10mL volumetric flasks containing 5mL of 0.2M HCl, 1 mL of stock nucleoside solution (10^{-4} M) and diluted with deionized water. Sample analysis was conducted by the alkaline quench method, adapted from Garrett and Mehta.^[37] All solutions, vials and flasks were equilibrated at the corresponding reaction temperature for 30 mins and then rapidly mixed in the volumetric flasks followed by transfer to 5mL reaction vials and placed in a temperature controlled ($\pm 0.1^{\circ}\text{C}$) water/ethylene glycol bath. At designated time points, 250 μL aliquots of the reaction mixture were removed and quenched in a UV-Vis cuvette with 250 μL of cold 1M NaOH solution. Full absorbance spectra were recorded and rates were determined by monitoring a decrease in absorbance at the largest wavelength change between the nucleoside and nucleobase. Kinetic experiments were conducted in duplicate or triplicate. Comparison of the alkaline spectra of nucleosides and the corresponding authentic nucleobases (obtained from commercial sources) were conducted to verify that the reaction spectra were indeed monitoring deglycosylation kinetics. For three nucleosides (deoxyisoguanosine, deoxyzebularine and deoxyxanthosine), deglycosylation rates were monitored directly by the UV-Vis spectrophotometer under the reactions conditions. The same setup was employed (pre-equilibration and mixing) but once mixed, a portion of the solution was transferred to a pre-equilibrated cuvette and completely filled with the reaction mixture and tightly closed with a well suited stopper to prevent evaporation. This was quickly placed in a cell holder that was specially designed to be controlled by a similar temperature controlled ethylene-glycol/water bath ($\pm 0.1^{\circ}\text{C}$) and kinetics were obtained by monitoring decrease in absorbance at one wavelength as in the other nucleosides. For reactions conducted at pH 2 and 3, (50°C), buffers of 0.1M sodium phosphate pH 2.0–2.1, 3.0–3.1 were used as dilution media and the experiments were conducted using the alkaline quench method and reaction progress was monitored spectrophotometrically as described above.

Ribonucleosides

The ribonucleoside reaction mixtures were prepared by using 10 mL volumetric flasks containing 5 mL of 0.2M HCl, 0.5–1 mL of stock nucleoside solution (10^{-4} M) and diluted with deionized water at room temperature, Aliquots of 600 μL were then used to completely fill reaction vials (0.6 mL, borosilicate glass, ChemGlass # CV-1510-0740), which were crimp sealed (ChemGlass # CV-3450-0008). Vials (about 10–13) were placed in a dry bath

(Grant Instruments # QBH2) and modified with sand and oil to ensure equal and consistent heating at the desired temperature. In all cases, it took vials less than 1 min to reach the desired reaction temperature with no noticeable impact on the reaction rate. At designated time points, a reaction vial was removed from the heat bath and submerged in ice-water. Upon cooling, 250 μ L aliquots of the reaction mixture were removed and quenched in a UV-Vis cuvette with 250 μ L of 1M NaOH solution. Full absorbance spectra were recorded and rates were determined by monitoring a decrease in absorbance at the largest wavelength change between the nucleoside and base. Kinetic experiments were conducted in duplicate and some cases triplicate.

RESULTS AND DISCUSSION

Kinetic and thermodynamic analyses

Rates of acid-catalyzed deglycosylation reactions were studied as function of pH and temperature and followed spectrophotometrically using an alkaline quench method (see Experimental). An example of spectral changes that occur over the time course of the deglycosylation reaction is shown in Figure 2 for 2,6-diaminopurine-2'-deoxyriboside. Also shown is the resulting pseudo-first order kinetic curve from the corresponding plot of a single-wavelength absorbance versus time, which acid-catalyzed deglycosylation reactions are known to exhibit.^[18] Additional examples of UV-Vis spectra of nucleoside deglycosylation reactions and their kinetic curves are included in the Supporting Information (Figures S1–S13).

Figure 3 shows compiled Arrhenius curves, which contain the extrapolated rate constants to 37°C for all the ribonucleosides and all but two of the deoxynucleosides studied (see caption Figure 3). These data along with the corresponding thermodynamic parameters of activation at 37°C, obtained from Eyring plots (see SI, Fig. S14), are summarized in Table 2 and arranged by descending deglycosylation rates.

The difference in the Gibbs free energy of activation, $G^\ddagger$, for acid catalyzed *N*-glycosyl cleavage at 37°C between a ribose and deoxyribose nucleoside for each of the bases studied, along with the corresponding rate enhancements ($k_{N\text{-deoxyribose}}/k_{N\text{-ribose}}$) are highlighted in Table 3. We include both kcal/mol and kJ/mol for $G^\ddagger$ (Table 2) and $G^\ddagger$ values (Table 3), as the latter unit appears a more appropriate metric for comparing smaller changes in $G^\ddagger$ values.

Figure 4 graphically depicts the results of deglycosylation kinetics of all the deoxynucleosides included in this study, determined at 50°C and pH ranging from 1 to 3. A tabular form of the same data obtained is given in Table 4. This specific temperature was selected as a compromise, to facilitate kinetic measurements as the reactions slow down with increasing pH, while at the same time represent conditions, which closely mirror the relative reactivity at 37°C.

Methodology and general observations on deglycosylation rates

The well-known stability of riboglycosidic linkages makes it impractical, in terms of time scale, to study deglycosylation kinetics at neutral pH and temperatures below approximately

100 °C. Depending on the nature of the heterocycle, carrying out these deglycosylation reactions at neutral pH and necessarily higher temperatures (usually > 120 °C) could complicate the kinetics of specifically monitoring *N*-glycosyl cleavage by accessing additional reaction pathways that lead to heterocyclic and monosaccharide degradation.^[3,16,19,38] Even with our method, the ribonucleoside deglycosylation kinetics had to be determined at higher temperatures (up to 85 °C) and extrapolate to 37°C for direct comparison to the deoxynucleosides. It is evident both from inspection of Table 2 or Figure 3 that the ribose moiety imparts significant stabilization to the *N*-glycosyl bonds compared to the deoxyribose series. Xanthosine, which displayed the fastest deglycosylation rate from the ribonucleoside series, still only proceeded at a rate that was a small fraction (3–4%) of the deglycosylation rates observed for deoxyguanosine or deoxyadenosine.

Deglycosylation rates under acidic conditions and heterocycles' basicity

The impact that the heterocycles' pK_a values have on base pairing and duplex stability has been recently discussed as a contributory pressure during the course of nucleic acids evolution and base selection.^[39] This is an intriguing notion since these pK_a values are unique to each heterocycle and may offer a seemingly direct explanation for why certain bases were selected and others not. In the deglycosylation reactions explored here, the acid/base properties of the nucleobases may also have a direct role.^[40] The ease of hydrolytic rupture of native purine *N*-glycosyl bonds in comparison to the native pyrimidines under physiological conditions is thought to result from the greater susceptibility of acid or metal ion catalyzed deglycosylation exhibited by the purines.^[34] Evidence to support this comes from relative pK_a values, the sites of protonation and the observed pH dependence on deglycosylation rates.^[18,40] Acid catalysis lowers the enthalpic activation parameter ($H^\ddagger$) of deglycosylation as it enhances the leaving group ability of the purine therefore destabilizing the *N*-glycosyl bond.

Table 5 lists the known pK_a values and first site of protonation under the conditions used in this study,^[41] along with the corresponding half-life values and enthalpy activation parameters for the ribonucleoside and deoxyribonucleoside deglycosylation reactions, respectively. It might be expected that a more basic purine with a higher pK_a value would exhibit a faster reaction rate than a less basic heterocycle. The values listed in Table 5 highlight, however, that the rates appear to be more significantly affected by the site of protonation and the corresponding nature of heterocyclic substitution. For example, among the ribonucleosides, the most basic heterocycle, 2,6-diaminopurine ($pK_a = 5.09$), was determined to have the longest $t_{1/2}$ value (8.8 days) and correspondingly highest enthalpic barrier ($H^\ddagger = 28.4$ kcal/mol), whereas the least basic nucleobase xanthine (pK_a range 0–0.8), exhibited the shortest ($t_{1/2} = 7.1$ hours) value and lowest barrier ($H^\ddagger = 25.2$ kcal/mol). This observation can be rationalized by the expected first sites of protonation for each heterocycle (N3 for 2,6-diaminopurine and N7 for xanthine) in addition to the different substituents on each. Protonation of the N7 destabilizes the *N*-glycosyl bond much more than protonation of the purine's N1 or N3.^[40] Furthermore, the electron withdrawing keto groups on Xan further enhances the leaving group ability whereas the electron donating amino groups on DAP decreases it. These observations can be extended to the other purine nucleobases. Isoguanine ($pK_a = 4.51$) and adenine ($pK_a = 4.15$) display similar pK_a values

but the first sites of protonation for each heterocycle are different as isoG is protonated at the N7 and A at the N3.^[41] Thus isoguanosine had a lower enthalpic barrier ($H^\ddagger = 25.7$ kcal/mol) in comparison to adenosine ($H^\ddagger = 26.1$ kcal/mol) and was observed to deglycosylate about 4.6 times faster than the latter. The first sites of protonation are the same for guanosine and inosine, both on N7, but it could be that the slower reaction rate observed for guanosine may be attributed to the presence of the electron donating amino group. Yet there remains the curious observation that the determined enthalpic barriers did not appear to correlate with the observed rate difference. Inosine contains a heterocycle that is much less basic ($pK_a = 1.9$) in comparison to guanosine ($pK_a = 3.30$) and was, expectantly, found to exhibit a $H^\ddagger$ with 2 kcal/mol value greater than guanosine. However, even with this enthalpic barrier difference, the reaction rate of inosine was actually 1.6-fold faster than guanosine. In this case, the entropic contribution ($S^\ddagger$) appears to become more significant in the overall observed rates (see Table 2). Related observations were found among the deoxynucleoside $t_{1/2}$ values that are also shown in Table 5.

As with the purines, the deglycosylation pathways of the pyrimidines are also known to proceed through an S_N1 type mechanism that involves the expulsion of the heterocycle leaving group followed by hydrolysis of the ribose oxocarbenium intermediate.^[34,40] While it is known that native pyrimidines are less susceptible than purines to acid catalyzed deglycosylation,^[18,34,42] not all pyrimidines are preordained for such stability. The alternative pyrimidine nucleosides, Zeb and dZeb have previously been shown to be susceptible to deglycosylation.^[11,43] To place the reactivity of dZeb/Zeb in perspective, we include here published data on related deoxypyrimidines that were conducted under the same reaction conditions (Figure 5, dC, and disoC). First, comparing the hydrolysis kinetics of dZeb/Zeb to Shaprio's study on deoxycytidine hydrolysis under similar conditions highlights the remarkable stability of the dC's *N*-glycosyl bond.^{[14],[44]} While the presence of the amino group on cytosine makes the pK_a value at the N3 position higher ($pK_a = 4.4$) compared to the N3 on 2-pyrimidinone ($pK_a = 2.1$), it also appears to significantly diminish its leaving group ability. The corresponding deglycosylation rate of dC at pH 1 and 37°C is $k = 1.7 \times 10^{-8} \text{ s}^{-1}$, indicating deoxycytidine deglycosylates at only 0.002 % the rate of dZeb and 0.4 % the rate of Zeb (Table 6). The reaction rate disparities can be explained by comparing the enthalpic barrier for dC hydrolysis under these conditions, which was calculated to be 37.2 kcal/mol. This is a full 13 kcal/mol greater than the value obtained for dZeb ($H^\ddagger = 24.1$ kcal/mol) and still 9 kcal/mol larger than the ribonucleoside Zeb.

Yet an even more dramatic difference in the relative stability of a pyrimidine *N*-glycosyl bond can result from a simple isomerization of the cytosine face. It was previously reported that under the same acidic conditions studied here (0.1M HCl),^[45] but at 40°C, disoC deglycosylates with a $t_{1/2} = 3.5$ mins. By comparison, dZeb deglycosylates at 40°C with a $t_{1/2} = 9.4$ mins and the determined values for dC at 40°C, gives a $t_{1/2} = 260$ days (Table 6). Thus, even with the exocyclic amino group, disoC still deglycosylates about 3 fold faster than dZeb, and over 100,000 times faster than its native isomer at 40°C (Table 6). It is still unknown why such a reaction rate disparity exists between the two isomeric nucleosides^[18] but it has been reported that the isocytosine heterocycle is also able to populate diverse tautomers under varying conditions.^[46, 47] We postulate that under the reaction conditions

reported^[45, 46] the formation of an alternative tautomer(s) of disoC might be contributing to its superior leaving group ability. Whatever mechanistic details eventually emerge, it is important to mention that while the sugar moiety is often the dominant contributor to *N*-glycosyl stability, the comparison of dZeb, Zeb, and disoC to dC clearly demonstrates that the heterocycles can significantly impact the deglycosylation reaction rates.

Rate enhancements determined from ribonucleoside and deoxyribonucleoside kinetics

It is well established that *N*-glycosyl stability drops as a nucleoside transitions from a riboglycosidic bond to a deoxyribose one, which can be measured as rate enhancements ($k_{N\text{-deoxyribose}}/k_{N\text{-ribose}}$).^[18,48] We have previously hypothesized that the higher stability of riboglycosidic bonds imparted by the ribose moiety could have been an advantageous feature for the base diversity in an RNA world.^[1, 22] The transition to DNA^[49] and consequentially the global weakening to *N*-glycosidic bonds^[50] may have provided a unique selection pressure that relied on the intrinsic stability of the *N*-glycosidic bond. However, a subtlety that we gleaned from the literature is that rate enhancements might actually vary. Thus from a biochemical evolutionary perspective, it occurred to us that if a variation in rate enhancements accompanied this transition from genomic RNA to DNA, some *N*-glycosyl bonds might suffer a larger destabilization than others, creating yet another varied population of hydrolytic stabilities in the early DNA world.^[22]

With respect to non-native nucleosides, little has been done to assess the changes in deglycosylation rates under the same experimental conditions. Thus we began to quantify the deglycosylation rate enhancements between a ribonucleoside and the corresponding deoxyribonucleoside for each heterocycle as it relates to our hypothesis.^[22] The data listed in Table 3 or graphically presented in Figure 6 display that, under the same conditions, relative *N*-glycosyl stability between ribonucleosides and deoxyribonucleosides do indeed vary. Even with this somewhat limited set of nucleosides, there were noticeable variations in the differences in the $G^\ddagger$ ($G^\ddagger$) values. Excluding the observation that the pyrimidine nucleosides containing the heterocycle 2-pyrimidinone displayed the smallest rate enhancement, little correlation could be made between the nature of the nucleoside and the ranking of its respective $G^\ddagger$ value. For example, the rate enhancement of the nucleosides containing Dap (1.4×10^3) was different than that for the ones containing A (7.9×10^2), yet in comparing the rate enhancements for Hyp to Xan, they exhibited nearly equal values (7.3×10^2). Also, a comparison of the rate enhancements between the isomeric structures of isoG (7.5×10^2) and G (5.9×10^2) proved to be different. Reasons for these $G^\ddagger$ variations are still unclear, but it appears that the $G^\ddagger$ between a ribo-glycosidic and 2'-deoxyribo-glycosidic bond under the same conditions is dependent on the specific nature of the attached heterocycle and not just a fixed value resulting from the differences contributed by the sugar moieties.

Importantly, rate enhancements between the deoxy- and the corresponding ribo-nucleoside are also expected to differ depending on pH and temperature and remain unique to the identity of the individual purine and pyrimidine heterocycle. In a study where we directly measured the deglycosylation rates of xanthosine and deoxyxanthosine at pH 2, 37°C (SI, Table S1) we observed a rate enhancement of 1.7×10^3 , which is different than the observed

rate enhancement of 7.3×10^2 measured at pH 1 and 37°C (Table 3). Another heterocycle, 5-hydroxyuracil, was observed under acidic conditions (1 M HCl) to exhibit a rate enhancement of around 3.7×10^2 in comparing deglycosylation rate of the corresponding nucleosides, 5-hydroxyuridine and 5-hydroxy-2'-deoxyuridine, at 60°C.^[42] Wolfenden has recently shown that under neutral pH and extrapolation to 25°C, the weakening of the glycosidic bond in moving from adenosine to deoxyadenosine exhibited only a 30-fold difference in reaction rates,^[19] in comparison to the 7.9×10^2 fold we observed under pH 1 and 37°C (Table 3). It appears that the variation of $G^\ddagger$ values is not just a phenomenon specific to acidic conditions, but may also be identified under neutral pH conditions, but more direct evidence for this will need to be obtained.

Deglycosylation rates of deoxyribonucleosides across low pH values

We postulated that stability of the glycosidic bond may have been a more important feature in the early DNA world than the RNA world because the formation of DNA abasic sites (which follows a deglycosylation reaction) may have presented the greatest threat to genetic stability.^[22] Consequently a selection for the nucleobases that contributed the most stable glycosidic bonds in polymeric DNA would have also influenced the evolution of nucleobase composition in the genetic alphabet. This suggests then that the native deoxynucleosides, in comparison to other plausible alternatives, may exhibit the most hydrolytically robust *N*-glycosyl bonds, possibly even across a range of conditions. Our initial systematic investigation, measuring deglycosylation rates of modified, and alternative, deoxynucleosides at 50°C, indeed highlights the greater hydrolytic stability of the native nucleosides across low pH values (Table 4 and Figure 7).

The modified or alternative purines used in this study (Figure 1 Table 1), were typically found to exhibit faster deglycosylation rates in comparison to the native deoxynucleosides. Also, our inclusion of dZeb in this study highlights the similarity in deglycosylation kinetics between a non-native pyrimidine with non-native purine nucleosides. As the pH was increased, dZeb exhibited extremely fast reaction rates (15–20 times faster) than the native nucleosides, at pH 3, with only deoxyxanthosine exhibiting rates that were faster (Table 4, and Figure 7). The only notable exception, however, is with the alternative nucleoside dDap, as it appeared to exhibit a deglycosylation rate that approached the native purines (Figure 7). Given the higher basicity of Dap (pK_a 10.77) compared to adenine (pK_a = 9.8) ^[39, 41] as a qualitative assessment of its leaving group ability in the form of a monoanion under physiological conditions, it is predicted that dDap will likely exhibit even slower reaction rates than deoxyadenosine at neutral pH. More systematic studies will need to be conducted at pH values closer to neutrality to see how the trends for these and other non-native nucleosides hold.

CONCLUSIONS

The nature of the heterocycle in nucleoside deglycosylation reactions may be considered secondary to the role of the sugar moiety. It is, however, the specific identity of the heterocycle that dictates the relative stabilities of *N*-glycosyl bonds within a nucleoside class and this seemingly small detail may have been significant in the natural selection of the

nucleobase repertoire. We have shown that even rather simple heterocyclic deviations from the native bases can significantly impact the reaction kinetics of *N*-glycosyl rupture. The native purine nucleosides and expectantly the native pyrimidine nucleosides do appear to exhibit the slowest deglycosylation kinetics under these conditions. We have also demonstrated that heterocycles have a unique contribution in the observed variation of rate enhancements when comparing deglycosylation kinetics of ribo→deoxyribonucleosides. While there were likely many selection pressures during the course of chemical evolution^[1,32] our observations presented here also suggest that the selection of the genetic alphabet was aided by the hydrolytic fitness of *N*-glycosyl bonds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the NIH, via grant number GM 069773, and the National Science Foundation, (GK-12 Socrates Fellowship to ACR) for generous support. We also acknowledge our anonymous reviewers for their helpful comments on the manuscript.

References

1. Rios AC, Tor Y. Israel Journal of Chemistry. 2013; 53:469. [PubMed: 25284884]
2. Hud, Nicholas V.; Cafferty, Brian J.; Krishnamurthy, R.; Williams, Loren D. Chem Biol. 2013; 20:466. [PubMed: 23601635]
3. Levy M, Miller SL. Proc Natl Acad Sci USA. 1998; 95:7933. [PubMed: 9653118]
4. Piccirilli JA, Benner SA, Krauch T, Moroney SE. Nature. 1990; 343:33. [PubMed: 1688644]
5. Benner SA. Acc Chem Res. 2004; 37:784. [PubMed: 15491125]
6. Karri P, Punna V, Kim K, Krishnamurthy R. Angewandte Chemie-International Edition. 2013; 52:5840.
7. Orgel LE. Crit Rev Biochem Mol Biol. 2004; 39:99. [PubMed: 15217990]
8. Sutherland JD. Cold Spring Harbor Perspect Biol. 2010; 2:a005439.
9. Bean HD, Sheng Y, Collins JP, Anet FAL, Leszczynski J, Hud NV. J Am Chem Soc. 2007; 129:9556. [PubMed: 17629276]
10. Kolb VM, Dworkin JP, Miller SL. J Mol Evol. 1994; 38:549. [PubMed: 11539446]
11. Sheng Y, Bean HD, Mamajanov I, Hud NV, Leszczynski J. J Am Chem Soc. 2009; 131:16088. [PubMed: 19839590]
12. Chen MC, Cafferty BJ, Mamajanov I, Gállego I, Khanam J, Krishnamurthy R, Hud NV. J Am Chem Soc. 2013; 10.1021/ja410124v
13. House CH, Miller SL. Biochemistry. 1996; 35:315. [PubMed: 8555190]
14. Shapiro R, Danzig M. Biochemistry. 1972; 11:23. [PubMed: 5009434]
15. Shapiro R, Kang S. Biochemistry. 1969; 8:1806. [PubMed: 5785205]
16. Schroeder GK, Wolfenden R. Biochemistry. 2007; 46:13638. [PubMed: 17973496]
17. Seela F, Gabler B. Helvetica Chimica Acta. 1994; 77:622.
18. Kochetkov, NK. Organic chemistry of nucleic acids, Part B. Budovskii, EI.; Kochetkov, NK., editors. Plenum Press; New York: 1972. p. 425
19. Stockbridge RB, Schroeder GK, Wolfenden R. Bioorganic Chem. 2010; 38:224.
20. Friedberg, EC.; Walker, GC.; Siede, W.; Wood, RD.; Schultz, RA.; Ellenberger, T. DNA Repair and Mutagenesis. 2. ASM Press; Washington, D.C: 2006.

21. Shapiro, R. Chromosome Damage and Repair. Seeberg, E.; Kleppe, K., editors. Plenum Press; New York: 1981. p. 3
22. Rios AC, Tor Y. Astrobiology. 2012; 12:884. [PubMed: 22984873]
23. Szathmary E. Nat Rev Genet. 2003; 4:995. [PubMed: 14631359]
24. Benner SA, Ricardo A, Carrigan MA. Current Opinion in Chemical Biology. 2004; 8:672. [PubMed: 15556414]
25. Callahan MP, Smith KE, Cleaves HJ, Ruzicka J, Stern JC, Glavin DP, House CH, Dworkin JP. Proc Natl Acad Sci USA. 2011; 108:13995. [PubMed: 21836052]
26. Borquez E, Cleaves HJ, Lazcano A, Miller SL. Origins Life Evol Biospheres. 2005; 35:79.
27. Levy M, Miller SL, Oró J. J Mol Evol. 1999; 49:165. [PubMed: 10441668]
28. Cleaves, HJ.; Lazcano, A. Chemical Evolution II: From the Origins of Life to Modern Society. American Chemical Society; Washington DC: 2009. p. 17
29. Kirnos MD, Khudyakov IY, Alexandrushkina NI, Vanyushin BF. Nature. 1977; 270:369. [PubMed: 413053]
30. Barks HL, Buckley R, Grieves GA, Di Mauro E, Hud NV, Orlando TM. Chem Bio Chem. 2010; 11:1240.
31. Nishikura K. Annu Rev Biochem. 2010; 79:321. [PubMed: 20192758]
32. Limbach PA, Crain PF, McCloskey JA. Nucleic Acids Res. 1994; 22:2183. [PubMed: 7518580]
33. Berg, JM.; Tymoczko, JL.; Stryer, L. Biochemistry. 6. W.H. Freeman; New York: 2007.
34. Gates KS. Chem Res Toxicol. 2009; 22:1747. [PubMed: 19757819]
35. Cheng Q, Gu J, Compaan KR, Schaefer HF. Chemistry – A European Journal. 2012; 18:4877.
36. Nuevo M, Milam SN, Sandford SA. Astrobiology. 2012; 12:295. [PubMed: 22519971]
37. Garrett ER, Mehta PJ. J Am Chem Soc. 1972; 94:8532. [PubMed: 4638983]
38. Larralde R, Robertson MP, Miller SL. Proc Natl Acad Sci USA. 1995; 92:8158. [PubMed: 7667262]
39. Krishnamurthy R. Acc Chem Res. 2012; 45:2035. [PubMed: 22533519]
40. Berti PJ, McCann JAB. Chem Rev. 2006; 106:506. [PubMed: 16464017]
41. Dawson, RMC. Data for biochemical research. 3. Clarendon Press; Oxford: 1986.
42. Garrett ER, Seydel JK, Sharpen AJ. Journal of Organic Chemistry. 1966; 31:2219.
43. Iocono JA, Gildea B, McLaughlin LW. Tetrahedron Lett. 1990; 31:175.
44. Extrapolated rate constant and enthalpic parameter was obtained from constructin of the Arrhenius and Eyring curve curve from available data of Shapiro and Danzig, 1972 (ref 14)
45. Seela F, He Y. Helvetica Chimica Acta. 2000; 83:2527.
46. Dracinsky M, Jans P, Ahonen K, Budesinsky M, Switzer. Eur J Org Chem. 2011; 8:1544.
47. Roberts C, Bandaru R, Switzer C. J Am Chem Soc. 1997; 119:4640.
48. Dworkin JP, Miller SL. Carbohydrate Research. 2000; 329:359. [PubMed: 11117319]
49. Lazcano A, Guerrero R, Margulis L, Oró J. J Mol Evol. 1988; 27:283. [PubMed: 2464698]
50. Lindahl T. Nature. 1993; 362:709. [PubMed: 8469282]

Figure 1.

Heterocycles corresponding to the nucleosides used in this study, including: 2,6-diaminopurine (Dap), hypoxanthine (Hyp), adenine (A), guanine (G), isoguanine (isoG), xanthine (Xan), and 2-pyrimidinone (Pyr).

Figure 2.

Deglycosylation of 2,6-Diaminopurine-2'-deoxyriboside at pH 3 and 50°C, kinetically monitored by UV-Vis spectroscopy using the alkaline quench method. Rate constants were determined by plotting absorbance vs. time at wavelengths with the greatest absorbance difference (arrow).

Figure 3.

Arrhenius plots of deglycosylation kinetics, at pH 1, of ribonucleosides and deoxyribonucleosides. Extrapolated data of ribonucleosides and four deoxyribonucleosides to 37 °C are indicated with an asterisk. The nucleosides isoG and dZeb are the only two that have reaction rates experimentally determined (duplicate and triplicate trials respectively) at 37 °C.

Figure 4.

Comparative deglycosylation rates at 50°C of 2'-deoxynucleosides at pH 1, 2 and 3. The average values for all deoxynucleosides at pH 2 and 3 were determined from direct measurements (in triplicate). At pH 1, only the value for dZeb was determined directly at 50°C (in triplicate); the other values were extrapolated from the Arrhenius plots (Figure 3).

Figure 5.
Structures of deoxyzebularine (dZeb), deoxycytidine (dC) and deoxyisocytidine (disoC)

Figure 6.

Graphical representation of rate enhancements (and $G^\ddagger$ values) for each of the nucleosides studied (from Table 3).

Figure 7. Kinetic stability of nucleosides studied relative to 2'-deoxyxanthosine. Plotting the nucleoside kinetics data from Table 4 relative to the reaction rates of dXan highlights the distribution of the *N*-glycosyl bond stabilities explored in this study.

Table 1

Nucleosides studied, their prebiotic relevance and occurrence in nucleic acids

Heterocycle	Prebiotic relevance ^a	RNA Biology ^a	DNA Biology ^a
2,6-Diaminopurine	M, P-Low ^[25–28]	N	R, S (to A) ^[29]
Hypoxanthine	M, P-High ^[25, 28, 30]	S-High (to G), B ^[31–33]	D (from A) ^[34]
Isoguanine	P-Low ^[28]	R ^[35]	D (from A) ^[35]
Xanthine	M, P-Mod ^[25, 28]	B ^[33]	D (from G) ^[34]
2-pyrimidinone	P ^[9, 36]	N	N
Guanine	M, P-Mod ^[25, 27–28, 30]	L ^[33]	L ^[33]
Adenine	M, P-High ^[25, 28]	L ^[33]	L ^[33]

^a Key: **B**=biosynthetic precursor, **D**=arises from damage to native letter, **L**=native letter, **M**=identified in meteorites, **N**=not yet identified, **P**=observed in prebiotic experiments, **R**=rare occurrence, **S**=surrogate letter, **High**=high abundance, **Mod**=moderate abundance, **Low**=low abundance.

Rate constants, half-lives, and thermodynamic activation parameters for acid catalyzed deglycosylation reactions in 0.1M HCl at 37°C

Table 2

Nucleoside	Rate, $10^6 k, s^{-1}$	$t_{1/2}^a$	G [‡] kcal/mol (kJ/mol)	H [‡] kcal/mol	S [‡] cal/mol•K
2,6-Diaminopurineriboside (Dap)	0.91 ± 0.03	8.8 d	26.75 ± 0.02 (111.9)	28.4 ± 0.8	5.4 ± 0.2
Adenosine (A)	0.98 ± 0.05	8.2 d	26.71 ± 0.03 (111.7)	26.1 ± 0.9	-1.87 ± 0.1
Guanosine (G)	1.17 ± 0.04	6.9 d	26.60 ± 0.02 (111.3)	24.9 ± 0.6	-5.3 ± 0.2
Inosine (Ino)	1.92 ± 0.02	4.2 d	26.29 ± 0.01 (110.0)	27.1 ± 0.2	2.6 ± 0.1
IsoGuanosine (IsoG)	4.58 ± 0.04	1.7 d	25.76 ± 0.01 (107.7)	24.8 ± 0.2	-2.9 ± 0.1
Zebularine (Zeb)	4.81 ± 0.05	1.7 d	25.73 ± 0.01 (107.6)	27.8 ± 0.3	6.8 ± 0.1
Xanthosine (Xan)	26.9 ± 0.5	7.1 h	24.67 ± 0.01 (103.2)	25.2 ± 0.4	1.8 ± 0.1
Deoxyguanosine (dG)	712 ± 20	16 m	22.65 ± 0.02 (94.7)	21.7 ± 0.5	-3.2 ± 0.1
Deoxyadenosine (dA)	774 ± 10	15 m	22.59 ± 0.01 (94.5)	22.9 ± 0.3	1.1 ± 0.1
Deoxyzebularine (dZeb)	835 ± 10	13.8 m	22.55 ± 0.01 (94.3)	24.1 ± 0.9	5.0 ± 0.1
2,6-Diaminopurine-deoxyriboside (dDap)	1360 ± 20	8.4 m	22.25 ± 0.01 (93.1)	22.9 ± 0.2	2.1 ± 0.1
Deoxyinosine (dIno)	1400 ± 40	8.5 m	22.25 ± 0.02 (93.1)	24.4 ± 0.5	7.0 ± 0.2
Deoxyisoguanosine (disoG)	3470 ± 70	3.32 m	21.67 ± 0.01 (90.7)	20.7 ± 0.1	-2.9 ± 0.1
Deoxyxanthosine (dXan)	18700 ± 100	37 s	20.63 ± 0.01 (86.3)	20.9 ± 0.3	1.3 ± 0.1

^a **Key:** d = days, h = hours, m = mins, s = secs

Table 3

Rate enhancements and $G^\ddagger$ values for deglycosylation reactions at pH 1 and 37 °C.

Nucleobase	$k_{N-deoxyribose}/k_{N-ribose}$	$G^\ddagger$ kcal/mol (kJ/mol)
2,6-Diaminopurine (Dap→dDap)	1.4×10^3	4.50 ± 0.02 (18.8 ± 0.1)
Adenine (A→dA)	7.9×10^2	4.12 ± 0.03 (17.2 ± 0.1)
Isoguanine (IsoG→dIsoG)	7.5×10^2	4.09 ± 0.01 (17.1 ± 0.1)
Hypoxanthine (Ino→dIno)	7.3×10^2	4.04 ± 0.02 (16.8 ± 0.1)
Xanthine (Xan→dXan)	7.3×10^2	4.04 ± 0.01 (16.9 ± 0.1)
Guanine (G→dG)	5.9×10^2	4.01 ± 0.03 (16.6 ± 0.1)
2-Pyrimidinone (Zeb→dZeb)	1.7×10^2	3.18 ± 0.01 (13.3 ± 0.1)

Table 4
Rate constants and $t_{1/2}$ values of deglycosylation reactions for deoxyribonucleosides at 50°C

Nuc.	<u>pH1</u>		<u>pH2</u>		<u>pH3</u>	
	$10^6 k, s^{-1}$	$t_{1/2}^a$	$10^6 k, s^{-1}$	$t_{1/2}^a$	$10^6 k, s^{-1}$	$t_{1/2}^a$
dA	3600 ±50	3.2 m	110 ± 6	1.8 h	17 ±1	11 h
dG	3040 ±80	3.7m	140 ±10	1.4 h	24 ±1	7.7 h
dDap	6300 ±80	1.8 m	180 ±9	1.1h	27 ±1	7.0 h
dIho	6900 ±200	1.6 m	450 ±50	25 m	73 ±2	2.6 h
disoG	14000 ±70	0.8 m	440 ±30	26 m	63 ±2	3 h
dZeb	4100 ±70	2.8 m	1700 ±30	7 m	370 ±2	31 m
dXan	74000 ±400	9.3 s	4200 ±100	2.7 m	720 ±10	16 m

^a **Key:** p = days, h =hours, m = mins, s = secs

Table 5

Comparison of purine nucleobase p*K*_a values to the half-life values of deglycosylation reactions at pH 1 and 37 °C.

Nucleobase	p <i>K</i> _a of nucleobase (protonation site) ^{a,41,40}	<i>t</i> _{1/2} ^a (ribonucleoside)	H ⁺ kcal/mol (ribonucleoside)	<i>t</i> _{1/2} ^a (deoxyribonuc.)	H ⁺ kcal/mol (deoxyribonuc.)
2,6-Diaminopurine	5.09 (N3)	8.8 d	28.4 ± 0.8	8.4 m	22.9 ± 0.2
Isoguanine	4.51 (N7)	1.7 d	24.8 ± 0.2	3.3 m	20.7 ± 0.1
Adenine	4.1 (N3)	8.2 d	26.1 ± 0.9	15 m	22.9 ± 0.3
Guanine	3.3 (N7)	6.9 d	24.9 ± 0.6	16 m	21.7 ± 0.5
Hypoxanthine	1.9 (N7)	4.2 d	27.1 ± 0.2	8.5 m	24.4 ± 0.5
Xanthine	0 – 0.8 (N7)	7.1 h	25.2 ± 0.4	37 s	20.9 ± 0.3

^aKey: d = days, h = hours, m = mins, s = secs

Table 6Comparison of pyrimidine nucleobase pK_a values and nucleoside half- life values at 0.1 M HCl

Nucleoside	pK_a of nucleobase (site of protonation) ⁴¹	$t_{1/2}$ at 37°C ^a	$t_{1/2}$ at 40°C ^a	Ref
dC	4.4 (N3)	4.7×10^2 d	2.6×10^2 d	¹⁴
disoC	4.2 – 4.4 (N3)	----	3.5 m	⁴⁵
dZeb	2.2 – 3.1 (N3)	13.8 m	9.4 m	This study
Zeb	2.2 – 3.1(N3)	1.7 d		This study

^a**Key:** d = days, h =hours, m = mins, s = secs