PREBIOTIC SYNTHESIS OF PYRIMIDINE NUCLEOSIDES

A Thesis
Presented to
The Academic Faculty

by

James P. Collins

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
December 2005
(Submitted posthumously)

PREBIOTIC SYNTHESIS OF PYRIMIDINE NUCLEOSIDES

Approved by:

Dr. Nicholas V. Hud, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Howard M. Deutsch
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. David M. Collard
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Thomas Orlando
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: November 21, 2005

ACKNOWLEDGEMENTS

The National Aeronautics and Space Administration (NNG04GJ32G) is acknowledged for research support. The NASA Summer High School Apprenticeship Research Program (SHARP) is acknowledged for support of the author as a mentor to two students at Georgia Institute of Technology.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

δ	chemical shift
DNA	Deoxyribonucleic acid
dH_2O	distilled water
HMBC	Heteronuclear Multiple-Bond Coherence
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence
NMR	Nuclear Magnetic Resonance
ppm	parts per million
RNA	Ribonucleic acid
TEAA	Triethyl ammonium acetate

SUMMARY

The problem of forming a glycosidic bond between ribose and the free nucleoside bases to produce β-nucleosides under plausible prebiotic conditions is commonly referred to in origin of life research as "The Nucleoside Problem". The lack of a general solution to this problem currently represents one of the largest stumbling blocks to the RNA world hypothesis and many other theories regarding the origin of life. Over thirty years ago the purine nucleosides were successfully synthesized by drying the fully-formed bases and ribose together in the presence of divalent metal ion salts. However, glycosidic bond formation by the pyrimidine bases has never been achieved under similar reaction conditions. This thesis describes the first plausible prebiotic synthesis of a pyrimidine nucleoside, demonstrated with the pyrimidine base analogue 2-pyrimidinone. Information provided by nucleoside-formation reactions involving 2-pyrimidinone and related pyrimidine bases should provide valuable insights into the possible mechanism by which glycosidic bond formation was accomplished on the prebiotic Earth.

1. INTRODUCTION

1.1. The RNA World Hypothesis

All current scientific theories regarding the origin of life agree on one point: Life evolved from a relatively simple self-replicating molecular system (Deamer, 1994). However, in present-day life self-replication is accomplished on the cellular level. Within living organisms, nucleic acids are required for the synthesis of proteins, and protein enzymes are required for the synthesis of nucleic acids. The interdependency of proteins and nucleic acids for mutual replication presents a paradox concerning the origin of life: Which came first, proteins or nucleic acids?

The famous Miller-Urey experiment, in which a model prebiotic atmosphere was subjected to electrical discharge for several days, demonstrated that amino acids could have spontaneously formed on the early Earth (Miller, 1953). Subsequent experiments showed that simply heating amino acids under drying conditions (at 170ºC) can produce peptides with random amino acid sequences (Fox, 1958). These random peptides, or "thermal copolymers", were also found to associate into spherical particles in aqueous solution (Fox, 1959). Some researchers cited these results as evidence that proteins were the first biopolymers of life. However, without nucleic acids it remained unclear how sequence information would have been stored or transferred in a protein-first world.

During the late 1960s Leslie Orgel, Francis Crick and Carl Woese each independently proposed that RNA was more likely than polypeptides to be the first polymer for life (Woese, 1967; Crick, 1968; Orgel, 1968). A common inspiration for these proposals was the realization that information storage is absolutely essential for a self-replicating system to evolve, and that RNA is a more likely ancient polymer than the only other informational polymer of contemporary life (i.e. DNA). Without experiments

to provide additional support, these proposals received little attention for more than a decade.

In the early 1980s two discoveries revolutionized our view of contemporary life, as well as our views on life's origins. The laboratories of Tom Cech and Sidney Altman each discovered different RNA molecules that perform catalytic processes in living cells (Guerrier-Takada, 1983; Kruger, 1982). These catalytic RNA molecules were named "ribozymes", emphasizing the fact that these are enzyme-like molecules composed of RNA. Prior to the discovery of ribozymes it was generally accepted that only proteins catalyzed chemical reactions in living cells. With the discovery of ribozymes it seemed that the paradox concerning which came first, proteins or nucleic acids, was solved. Because RNA could both store information and catalyze chemical reactions, it had both of the basic features believed necessary for self-replication and evolution. This prompted Walter Gilbert in 1986 to coin the phrase "The RNA world" in reference to a time in early evolution when life used RNA for information storage and chemical catalysis (Gilbert, 1986). Later, Noller and co-workers would also demonstrate that the essential components of the ribosome (the protein-RNA machine responsible for protein synthesis) are RNA and that ribosomal proteins play a more supporting role (Noller, 1992). These results were taken as evidence that RNA "invented" protein synthesis, a scenario that fit perfectly with the RNA world hypothesis.

In 1990 the laboratories of Szostak and Gold invented techniques for the selection of RNA molecules with specific properties from pools of RNA molecules containing random sequences (Ellington, 1990; Tuerk, 1990). Using such techniques, a number of laboratories over the past fifteen years have produced an impressive array of RNA molecules with novel catalytic abilities (Joyce, 1999a). Many researchers have cited this work as evidence that RNA can accomplish almost any task that is now carried out in life by proteins, and thus provides firm support for the RNA world hypothesis.

Despite the general enthusiasm that exists among origin of life researchers for the RNA world hypothesis, major gaps remain in our knowledge between the abiotic chemistry capable of producing the small molecule building blocks of life (e.g. amino acids, sugars, nucleotide bases) and the processes that gave rise to the first RNA, or RNA-like, polymers. Plausible prebiotic reactions have been described for how the nucleotide bases of RNA and ribose could have formed on the prebiotic Earth (Joyce, 1999b), but how these molecules became part of the same polymer is a mystery. This problem is the motivation for the research described in this thesis.

1.2. The Nucleoside Problem

The problem of forming a glycosidic bond between ribose and the free nucleoside bases to produce β-nucleosides under plausible prebiotic conditions, commonly referred to in origin of life research as "The Nucleoside Problem", represents one of the most significant problems for the RNA world hypothesis and many other theories regarding the origin of life. A number of laboratories have searched in vain for a solution to how nucleosides could be formed from preexisting bases and ribose (Zubay, 2001). Early efforts looked promising, as Orgel and co-works demonstrated decades ago that adenine spontaneously couples to ribose upon heating and drying in the presence of divalent metals ions (Figure 1.1) (Fuller, 1972). However, β-adenosine is only produced in modest yields by this reaction, and the dominant species formed is ribose linked to the amine group of adenine (Zubay, 2001). This simple approach to the nucleoside problem proved to be even less promising with other bases, as conditions have not been found in which unmodified ribose forms any detectable amount of β-nucleoside (or any nucleoside) with guanine, cytosine or uracil. It is possible that the limited solubility of guanine is a major impediment to the formation of guanosine. The pyrimidine bases, on the other hand, are considerably more soluble than adenine in aqueous solution. Thus, it appears that the reactivity of the pyrimidine bases with ribose is fundamentally less than that of adenine.

It has been proposed that the delocalization of charge on N1 of the pyrimidine bases, which must initiate a nucleophilic attack toward the aldehydic carbon of ribose, makes pyrimidine glycosidic bond formation difficult (Ingar, 2003). Therefore, one is lead to conclude that the chemistry of pyrimidine nucleoside formation transpired by a very different reaction pathway than the simple coupling of the pyrimidine bases and ribose under drying conditions, *or* that different pyrimidine bases preceded those that are now found in RNA.

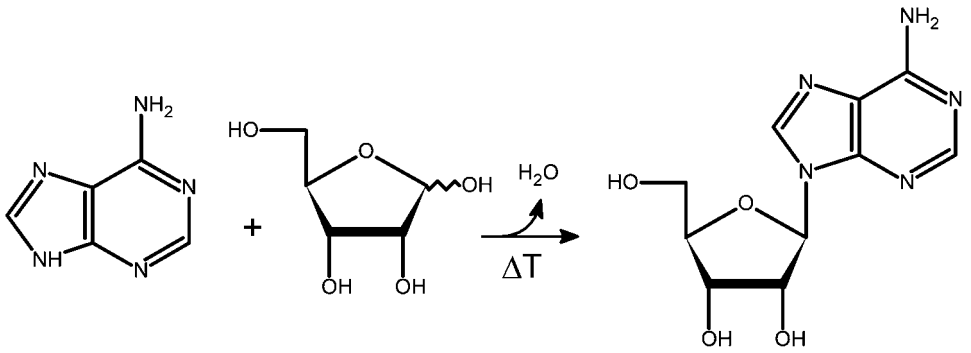

Figure 1.1 Synthesis of β-adenosine by heating and drying adenine with ribose.

1.3. Alternative Approaches to the Nucleoside Problem

Some researchers have attempted to prepare pyrimidine nucleosides by building the pyrimidine base directly on ribose in what could be argued to be a plausible prebiotic reaction. The stepwise synthesis of the base cytosine on a preformed sugar was accomplished by Sanchez and Orgel in 1970 (Sanchez, 1970). This reaction proceeds by the addition of cyanamide to the C1' and C2' hydroxyls of a closed pentose sugar, yielding a bicyclic product called an aminooxazoline derivative (Figure 1.2). This product then reacts with cyanoacetylene to give α-cytidine and β-arabinosylcytidine when the pentose sugars used are ribose and arabinose, respectively (the reaction with ribose is shown in Figure 1.2). These reactions are very regioselective, yielding very little

β-cytidine and α-arabinosylcytidine, respectively. The cytosine nucleosides can then be hydrolyzed to form the uridine nucleosides, thereby providing an overall pathway for the formation of both pyrimidine nucleosides. These experiments have been revisited and expanded in recent years by the Sutherland laboratory as part of an attempt to understand the mechanism and products of this reaction, and to explore similar approaches to step-wise purine synthesis on sugars (Saewan, 2005; Ingar, 2003). Sutherland and co-workers have been able to prepare β-cytidine by a reaction similar to that of Sanchez and Orgel. However, the Orgel and Sutherland syntheses require the use of a phosphorylated ribose, which is of questionable prebiotic relevance (Keefe, 1995).

There are some virtues to the approach of building a base on a sugar in regards to prebiotic chemistry. For example, the addition of a primary amine to C1' of the sugar allows the sugar to cyclize through a stable aminal bond, making the sugar resistant to degradation, as the rate of sugar degradation is directly proportional to the amount of time the sugar spends in the open-chain aldehyde form (Larralde, 1995). Additionally, sugars have built-in reactive groups (i.e. hydroxyls) that strongly influence the stereochemistry of products, as well as the ring structure of the sugar (e.g. furanose is favored over pyranose for ribose and arabinose, pyranose is favored for xylose) (Saewan, 2005). However, it is not clear that these advantages outweigh the disadvantages, such as the fact that ribose yields α-nucleosides. Perhaps the most obvious problem with the route used by Sanchez and Orgel to build nucleosides on a preexisting sugar, in regards to the RNA world, is that an analogous synthesis has not been found for production of the purine nucleosides.

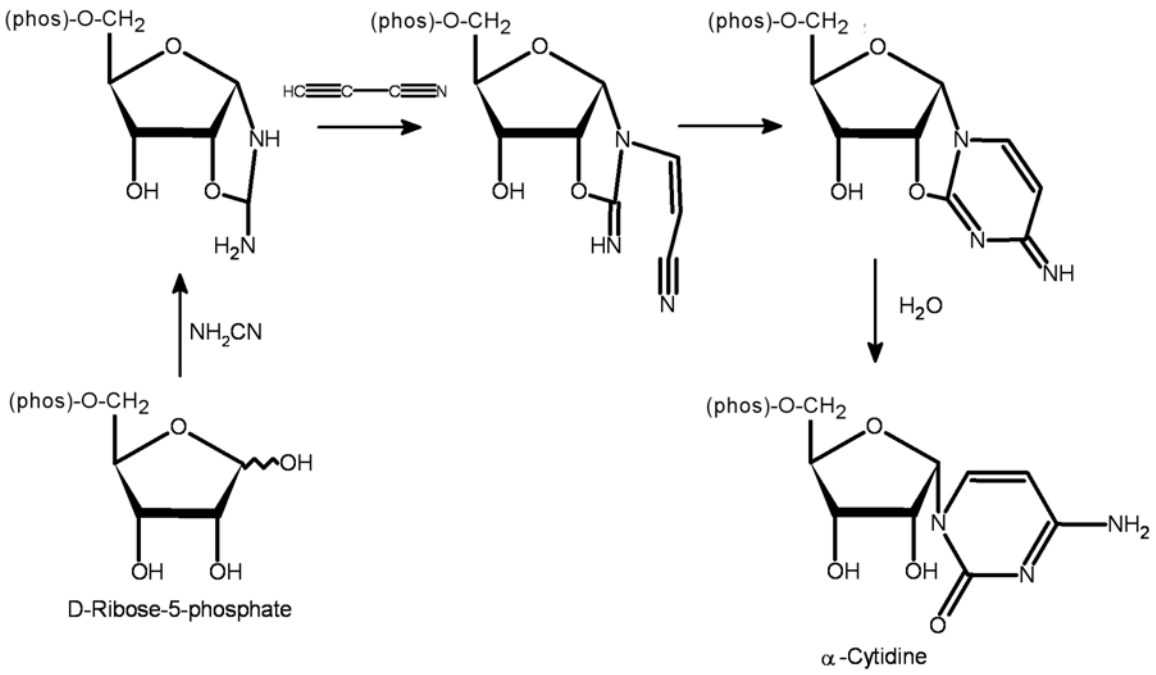

Figure 1.2 Synthesis of α-cytidine by the method reported by Sanchez and Orgel, 1970.

Still one more alternative hypothesis for the original abiotic synthesis of the nucleosides has been put forth by Saladino et al., in which it is proposed that sugars (or acyclic carbohydrates) were synthesized on preexisting bases. In a one-pot synthesis, catalyzed by TiO_2, Saladino et al. have demonstrated the production of purines and pyrimidines from formamide in reasonably good yield (Saladino, 2003). In addition to the synthesis of all the bases (with the exception of guanine/hypoxanthine), formyl purines and acyclonucleosides were also detected in the product mixture. The production of carbohydrate chains arose due to the presence of formaldehyde in the reaction mixture, generated by the photooxidation of formamide in the presence of TiO_2. The carbohydrate chain was only observed to grow to three carbons in length, however, so long as the polymerization rate of formaldehyde exceeds the rate of sugar degradation, the length of these chains should be able to reach 4, 5, or 6 carbons – lengths corresponding to

previously synthesized non-natural nucleic acid backbones containing tetroses, pentoses, and hexoses (Eschenmoser, 1999).

Saladino et al. recognized that their formylpurine products suggest the possibility of building a sugar stepwise from the formylpurine derivatives, and that this approach is a possible alternative route to nucleoside synthesis, an alternative in which the glycosidic bond is formed before the sugar (Saladino, 2003). There are two main problems with this idea for pyrimidine nucleoside synthesis. First, only purine acyclonucleosides were detected in the product mixture of Saladino et al., suggesting that their prebiotic synthesis is subjected to the same problems as those of Fuller et al. Second, the acyclonucleoside synthesized by Saladino et al. will not be able to cyclize without first reducing the carbonyl to an alcohol. When this step is performed, however, a hemi-aminal would be formed at C1' of the sugar, which is then exchangeable and no different than forming a glycosidic bond from a free sugar and a free base. However, if the reduction of the C1' carbonyl only occurred when water activity was very low, the exchange of the proto-glycosidic bond would not be as favorable.

1.4. Current Status of the Nucleoside Problem

As discussed above, the question of how nucleosides were first synthesized on the prebiotic Earth is still far from being resolved. For the plausible prebiotic reactions that have successfully produced nucleosides, not one of these reactions is able to produce all four nucleosides, or even one purine and one pyrimidine. The nucleoside problem clearly remains one of the most troublesome stumbling blocks to the RNA world hypothesis. Because the chemistry that resulted in the original abiotic synthesis of the nucleosides is not obvious, it is likely that the eventual solution to this problem will provide much information regarding the molecules of early life on Earth, such as which bases and sugars were ancestors to those that exist today in RNA and DNA.

2. A NEW MODEL SYSTEM FOR STUDYING PREBIOTIC NUCLEOSIDE SYNTHESIS

Over thirty years ago Orgel and co-workers demonstrated the prebiotic synthesis of adenosine by drying adenine and ribose from a solution containing $MgCl_2$ and $MgSO_4$ (Fuller, 1972). However, their attempts to synthesize cytidine and uridine under the same reaction conditions demonstrated that glycosidic bond formation for pyrimidine nucleosides does not occur as readily as it does for purines (Fuller, 1972). The common belief for the discrepancy between purine and pyrimidine condensation with sugars is that the lone pair of electrons on the purine N9 initiates a nucleophilic attack on the aldehydic carbon of ribose. This reaction is not likely to occur for the pyrimidine bases because the tautomerization of uracil, thymine, and cytosine delocalizes the electron density away from N1 (Figures 2.1-2.3), creating a kinetic barrier to glycosidic bond formation (Ingar, 2003). However, life may have started with different pyrimidine bases that more easily formed nucleosides, and these bases could have been replaced after an enzymatic mechanism for nucleoside synthesis had evolved.

There are pyrimidine base analogues that bear the appropriate arrangements of substituents on the heterocycle ring to increase or maintain electron density on a ring nitrogen (e.g. 2-pyrimidinone). These pyrimidine base analogues would thus be able to initiate nucleophilic attack of the sugar aldehyde under reaction conditions that have been demonstrated for purine nucleoside synthesis. Once the glycosidic bond is formed, these pyrimidine nucleoside analogues could then have been chemically modified to the natural pyrimidine nucleosides in the prebiotic environment. Alternatively, the pyrimidine analogues could have functioned in base pairing with natural or non-natural purines until catalytic polymers evolved with the capability of synthesizing cytidine, uradine, or thymidine, which then replaced the pyrimidine analogues with the natural nucleosides.

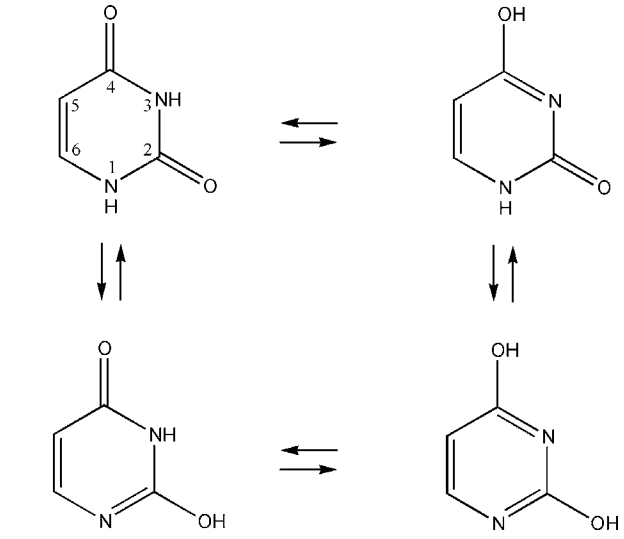

Figure 2.1 Uracil tautomers

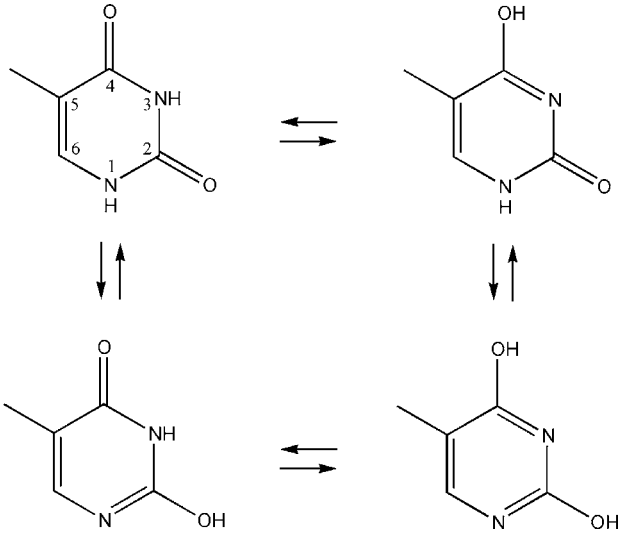

Figure 2.2 Thymine tautomers

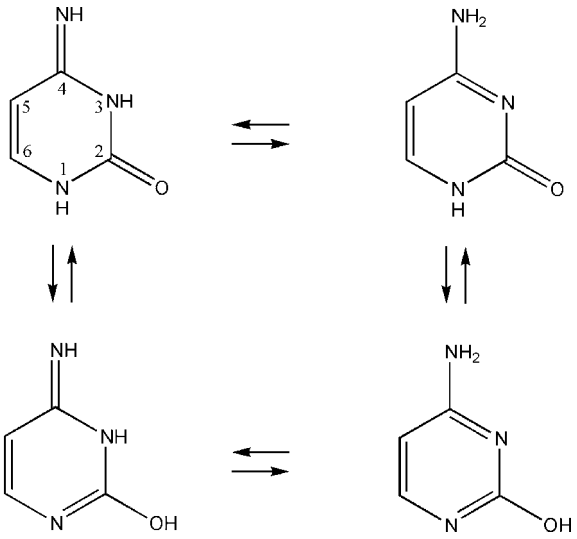

Figure 2.3 Cytosine tautomers

There are several commercially available pyrimidine bases that could be used to test the requirement for localization of electron density on N1 to effect condensation of the base and sugar. The base 2-pyrimidinone (Figure 2.4) was chosen as the model system and the best candidate for pyrimidine nucleoside formation for several reasons. First, it has only one electron-withdrawing group allowing the nitrogens to maintain more electron density than the natural pyrimidines. Second, the nitrogens are equivalent, making it possible to attach the base onto a sugar in only one position, unlike adenine, for instance, which was shown by Orgel et al. to attach to the sugar through N9, N7, and the primary amine group (Fuller, 1972). Reducing the number of synthesized 2-pyrimidinone

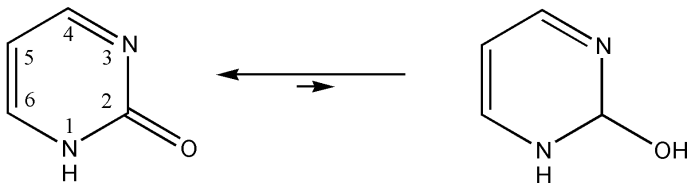

Figure 2.4 2-Pyrimidinone tautomers. The dominant form of the base in solution is the keto tautomer (Katritzky, 2000)

nucleosides not only makes the prebiotic chemistry more plausible as the distribution of products is reduced, but it also makes the analysis of the reaction products easier. Third, once the nucleoside is synthesized, it would be possible to modify the 2-pyrimidinone nucleoside into the cytosine nucleoside through amination, which can then spontaneously deaminate to uracil (Shapiro, 1966) (Figure 2.5), or oxidize 2-pyrimidinone into uracil, then aminate to cytosine (Figure 2.5). And last, the β-furanosyl 2-pyrimidone ribonucleoside, zebularine, is commercially available and can be used as a standard to verify 2-pyrimidinone nucleoside synthesis.

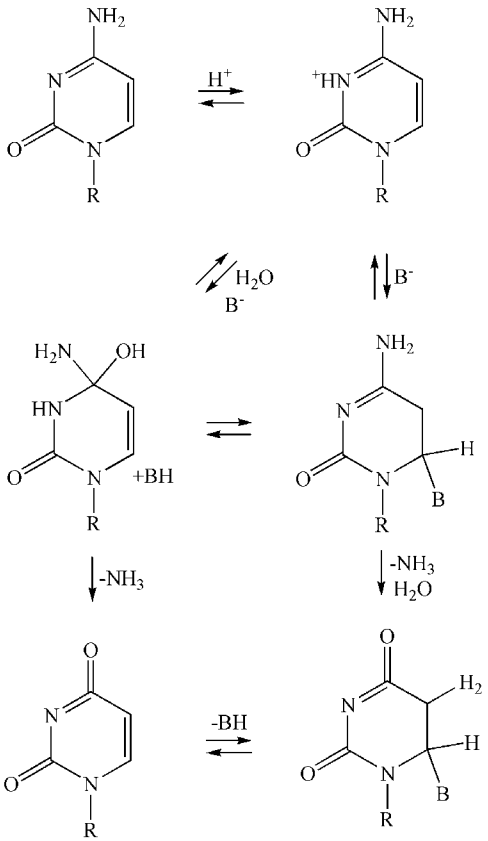

Figure 2.5 Conversion between cytosine and uracil (Shapiro, 1966).

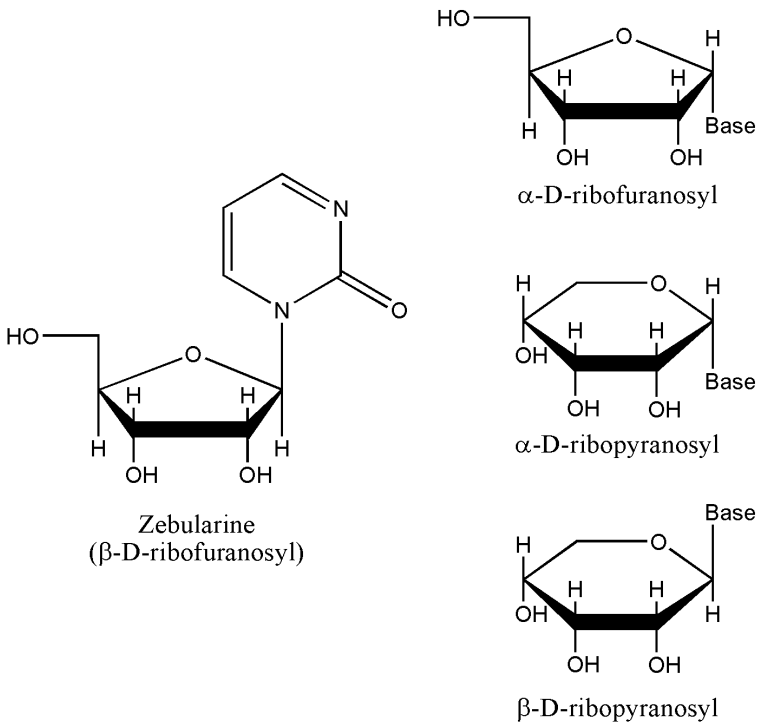

Figure 2.6 The four possible products of the condensation of 2-pyrimidinone and ribose. Zebularine, the β-D-ribofuranosyl nucleoside, most closely resembles the natural pyrimidine nucleosides. The 2-pyrimidinone base is abbreviated as "Base" in the other sugar anomers.

The prebiotic synthesis of the zebularine ribonucleoside from the condensation of 2-pyrimidinone and ribose can result in four possible products: α-furanosyl, β-furanosyl, α-pyranosyl, and β-pyranosyl ribonucleosides (Figure 2.6). The equilibrium distribution of ribose into its anomers suggests that the dominant product of the nucleoside synthesis would be the β-D-ribopyranosyl nucleoside, if anomers are equally reactive (Figure 2.6) (Carey, 1996). However, the reaction conditions under which the nucleosides are synthesized and the hydrogen bonding characteristics of the base can greatly influence the distribution of products (Kett, 1997; Fuller, 1972). Although the β-furanose form (zebularine) most closely resembles the extant pyrimidine ribonucleosides, RNA-like polymers have been synthesized containing pyranosyl forms of pentoses, and these analogue oligonucleotides have been shown capable of base pairing and forming double

helices (Bolli, 1997; Eschenmoser, 1999). Therefore, the synthesis of the β-D-ribofuranosyl zebularine may not have been critical for the incorporation of the 2-pyrimidinone nucleoside into prebiotic oligonucleotides. Together, the attributes of 2-pyrimidinone described here suggest that this pyrimidine and its associated nucleoside, zebularine, represent ideal prospects for the development of a new model system for prebiotic nucleoside synthesis.

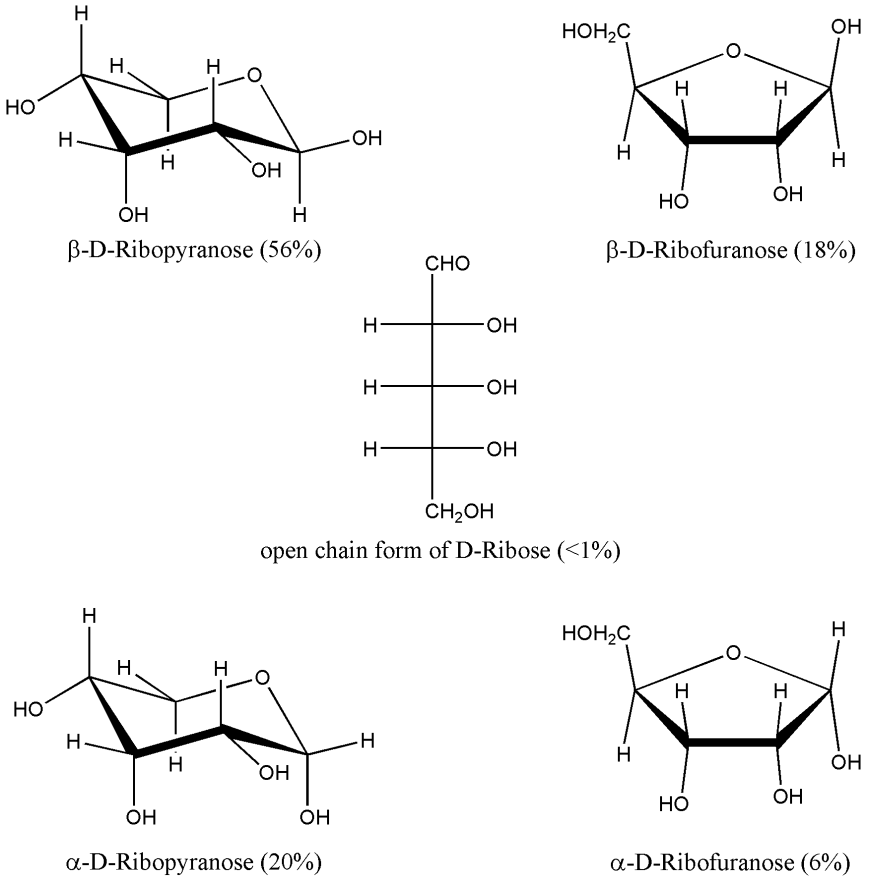

Figure 2.7 Equilibrium distributions of ribose in solution (Carey, 1996).

3. A PREBIOTIC SYNTHESIS FOR ZEBULARINE

3.1 Introduction

Many of the chemical reactions involved in the synthesis of oligonucleotides, carbohydrates, and polypeptides are condensation reactions, which release a water molecule with the formation of each bond (Schopf, 1999). The formation of the glycosidic bond is just such a reaction. Thus, the synthetic organic synthesis of a nucleoside from a free base and ribose would involve anhydrous solvents and scavenging the water molecules released during condensation in order to drive the reaction forward to completion. However, life very likely emerged in an aqueous environment in the pH range of 6-9 (Bengston, 1994). Therefore, a more prebiotically-plausible route toward condensation is to use dehydration by way of evaporative drying, which is described in this thesis for the synthesis of the 2-pyrimidinone nucleosides.

In this chapter reaction products resulting from the heating and drying of 2-pyrimidinone and ribose are described in detail. It is shown that a simple and plausible prebiotic reaction involving these two molecules results in the production of at least three pyrimidine nucleosides. One of these nucleosides is confirmed as the β-ribofuranosyl-2-pyrimidinone nucleoside, zebularine.

3.2 Materials and Methods

3.2.1 Synthesis of 2-pyrimidinone nucleosides

Each nucleoside reaction contained 5 mM 2-pyrimidinone base with 75 mM ribose in 100 μl dH_2O, pH 2.5. The samples were dried on a watch glass in an oven at 85˚C for 2 hr, and rehydrated in dH_2O. The 1:15 base to sugar ratio was chosen based on the plausible prebiotic synthesis reaction described for purine nucleosides (Fuller, 1972).

All reactions were performed in duplicate and products separated by HPLC. Purification of these products was achieved by the HPLC protocols below on a Phenomenex ODS semi-preparative column (10 mm x 250 mm, 5 μ).

3.2.2 HPLC Gradient Conditions

Reaction products were HPLC separated on a Phenomenex ODS analytical column (4.6 mm x 250 mm, 5 μ) with mobile phase conditions of 10-30% B (0-5 min), 30-50% B (5-30 min), flow rate = 1.0 ml/min. Mobile phase A was 0.1 M triethyl ammonium acetate (TEAA), and mobile phase B was 25% (v/v) acetonitrile, 75% (v/v) 0.1 M TEAA.

3.2.3 HPLC Isocratic Conditions

Reaction products were HPLC separated on a Phenomenex ODS analytical column (4.6 mm x 250 mm, 5 μ) with 0.1 M triethyl ammonium acetate as the mobile phase.

3.2.4 UV Spectrometry

UV-Vis absorbance measurements were performed in dH_2O using a HP 8453 UV-Vis diode array spectrophotometer.

3.2.5 NMR Spectroscopy

All NMR spectra were collected on a Bruker DRX500 in D_2O at 25ºC. Proton chemical shifts for the zebularine standard were assigned based on HSQC and HMBC ^{1}H – ^{13}C coherence spectroscopy (Appendix Figures A.1 – A.2). The ^{1}H assignments for the reaction products HPLC Peaks 1a, 2a, and 2b were made by comparison of their COSY spectra to the COSY spectrum of the zebularine standard.

3.3 Results

As discussed in Chapter 2, 2-pyrimidinone and D-ribose have the potential to condense into four different nucleosides: α-furanosyl, β-furanosyl (zebularine), α-pyranosyl, and β-pyranosyl configurations (Figure 2.6). To discern which of the four possible nucleoside products are synthesized, the reaction products were characterized by HPLC and NMR spectroscopy. HPLC analysis of the reaction products exhibited two primary peaks and two minor peaks with relative mobilities similar to zebularine, labeled as peaks 1a, 2a, 1b, and 2b, respectively (Figure 3.1).

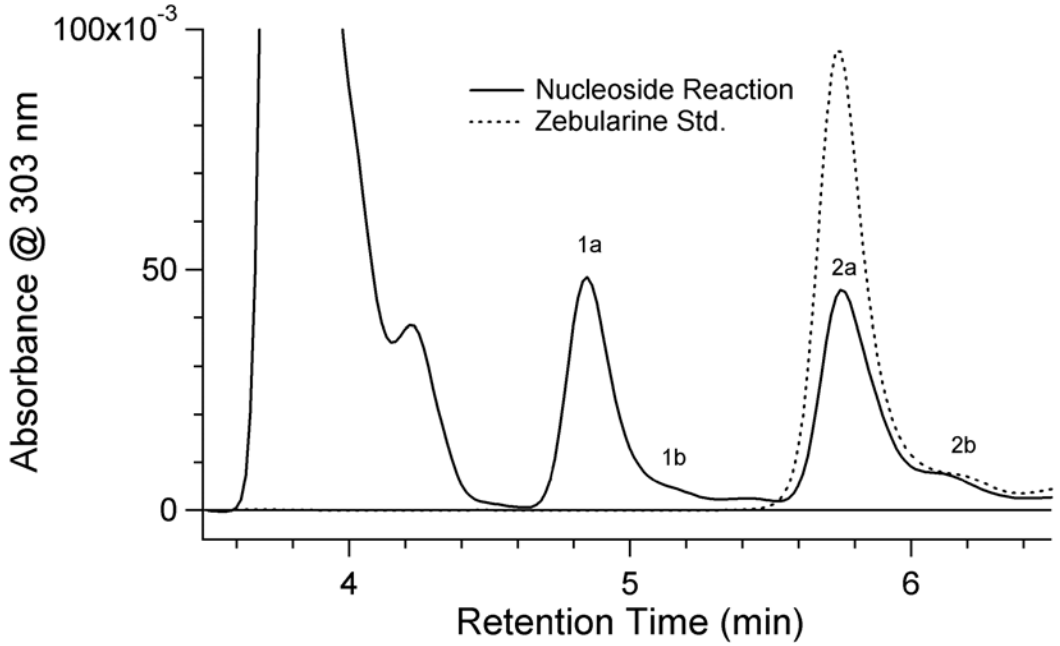

Figure 3.1 HPLC chromatographs illustrating the relative mobilities of the zebularine standard (dotted line) and four of the nucleoside reaction products (solid line) under gradient elution conditions (Materials and Methods).

Based upon the column mobilities of the zebularine standard and the reaction products under the gradient mobile phase conditions (Materials and Methods), it appeared

that HPLC Peak 2a was zebularine (i.e., the β-furanosyl nucleoside). [Note: Figure 3.1 shows that the zebularine standard also contains a small amount of the nucleoside conformer that is eluting as HPLC Peak 2b. This impurity can be seen by NMR (data not shown), and matches the NMR of Peak 2b in Figure 3.6C. Based upon NMR resonance peak integration, this impurity is <0.4% of the total nucleoside in the standard.] To further test the assignment of HPLC Peak 2a as zebularine, two more HPLC experiments were performed. First, the mixture of reaction products was spiked with an authentic zebularine standard, which showed that only HPLC Peak 2a increased in intensity, and without any apparent shoulders, further supporting that this reaction product is zebularine (Figure 3.2). Second, the nucleoside product mixture and the zebularine standard were both run under an isocratic mobile phase condition. If the zebularine and HPLC Peak 2a

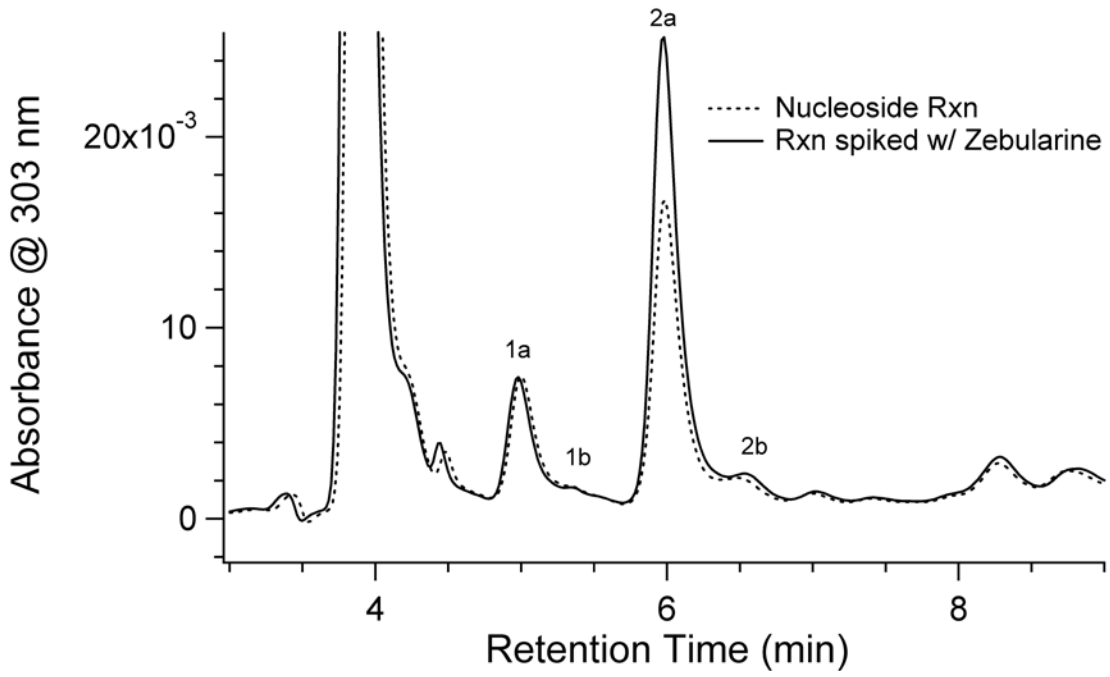

Figure 3.2 HPLC chromatographs of the nucleoside reaction products (dotted line) and the reaction products spiked with the zebularine standard (solid line). The coincident elution under the gradient elution conditions (Materials and Methods) of one of the nucleoside reaction products (Peak 2a) spiked with the zebularine nucleoside standard suggests that zebularine was synthesized in the nucleoside reaction.

only coincidentally exhibited the same column retention time in the gradient mobile phase, but were in fact two distinct molecular species, then it would be highly unlikely that two different molecules would also have the same mobility under the isocratic condition. The zebularine standard and HPLC Peak 2a were found to have the same mobility under the isocratic condition, providing additional support that HPLC Peak 2a is zebularine (Figure 3.3).

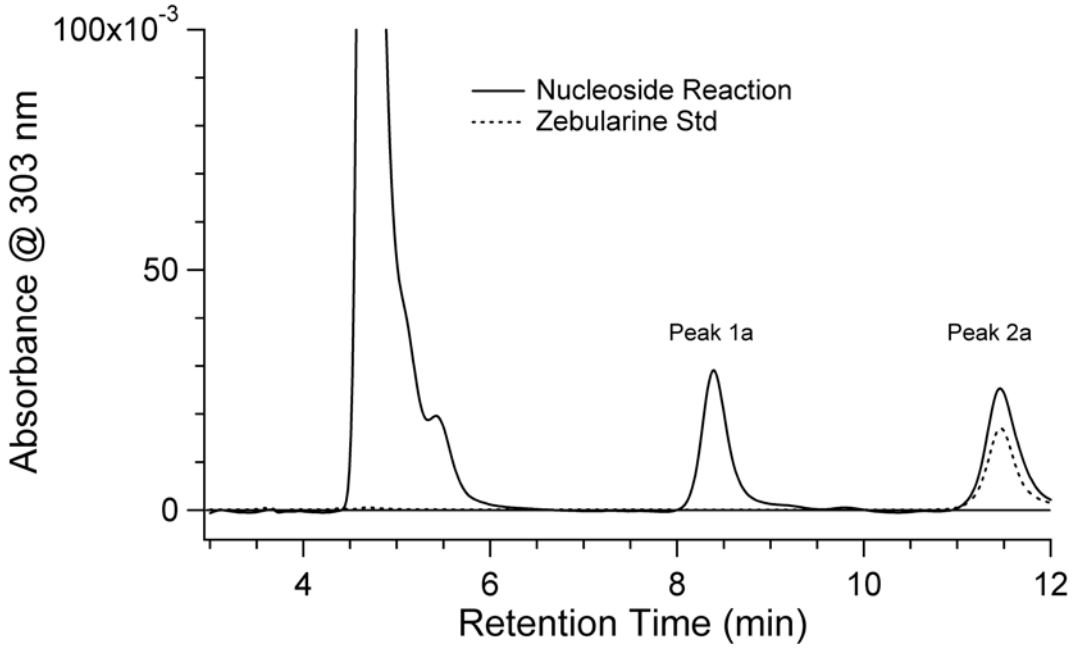

Figure 3.3 HPLC chromatographs illustrating the relative mobilities of the zebularine standard (dotted line) and the nucleoside reaction products (solid line) under the conditions of an isocratic elution (Materials and Methods).

Next, the suspected nucleoside reaction products were HPLC purified using the semi-preparative isocratic mobile phase conditions. The fractions with a UV spectrum that matched zebularine (Figure 3.4) were subsequently analyzed by NMR. The NMR of HPLC Peak 1a (isolated under the isocratic condition) is shown in Figure 3.5A. The presence of an aromatic compound is clear by the protons with chemical shifts in the

range of 6 – 8 ppm. Resonance splitting patterns and peak integration are consistent with HPLC Peak 1a being a 2-pyrimidinone nucleoside. However, in the region of the sugar protons (3 – 6 ppm), there are many extraneous peaks from impurities that prevent the identification of the sugar protons of the nucleoside (Figure 3.5A). These impurities must be compounds that co-elute with HPLC Peak 1a, but that did not absorb at 303 nm, the wavelength observed for the HPLC analysis. To remove these impurities, the NMR sample for HPLC Peak 1a was subjected to a second round of HPLC purification using the semi-preparative gradient mobile phase conditions. The use of the gradient removed the majority of the impurities from the nucleoside of interest, as seen in Figure 3.5B. This procedure was also carried out for HPLC Peak 2a to obtain a high purity sample for NMR analysis.

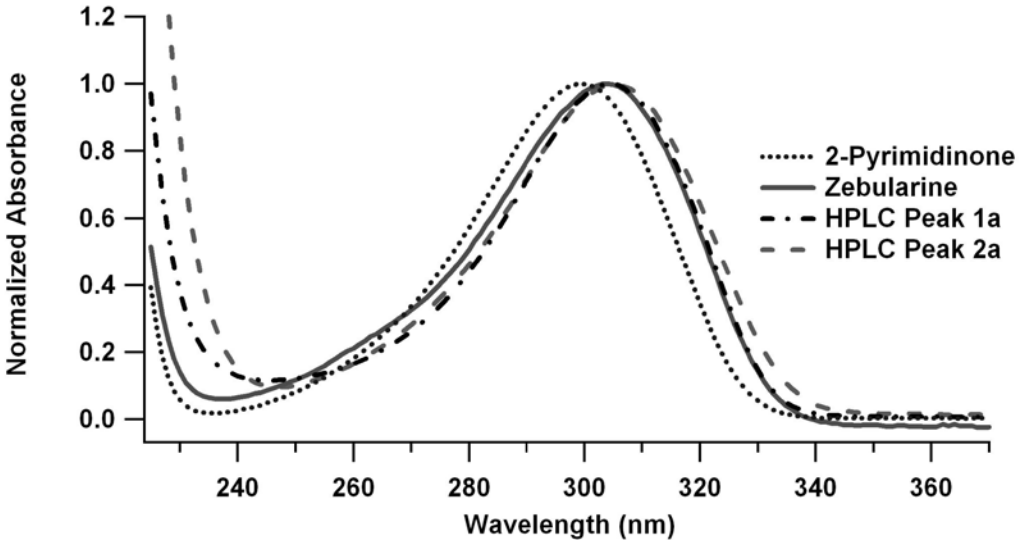

Figure 3.4 Normalized UV Spectra of the 2-pyrimidinone base standard, the zebularine standard, and HPLC Peaks 1a and 2a. The maximum absorbance for the free base (dotted line) occurs at 299 nm whereas, the maxima for zebularine (solid line), HPLC Peak 1a (dot-dashed line), and HPLC Peak 2a (dashed line) are at 304 nm.

The NMR spectra of the 2-pyrimidinone base, HPLC Peak 1a, HPLC Peaks 2a and 2b, and the zebularine standard are shown in Figure 3.6A-D, respectively. HPLC Peaks 2a and 2b were collected in the same HPLC fraction and are seen as the major and minor species, respectively, in the NMR spectrum of Figure 3.6C. The spectrum of the base (A) shows that H4 and H6 are equivalent through rapid tautomerization, where protons H4 and H6 possess the same chemical shift, and H5 is coupled to two protons (H4 and H6) with identical coupling constants. However, once the 2-pyrimidinone base is attached to a sugar, as in the zebularine standard, protons H4 and H6 become non-equivalent (Figure 3.6D). Comparison of the spectra for zebularine to the major species in Figure 3.6C shows the same peaks in both samples, further proof that β-ribofuranosyl nucleoside, zebularine, was synthesized. Additionally, the COSY spectrum of HPLC Peaks 2a/2b shows the same correlations between base protons and between sugar protons as those for zebularine (Figures 3.7 – 3.9). Although the NMR spectra of HPLC Peaks 1a and 2b are distinct from the zebularine spectrum in chemical shifts and coupling constants, each of the base and sugar protons of HPLC Peak 1a and 2b can be correlated resonance-by-resonance with the zebularine standard, demonstrating that these two species are, indeed, 2-pyrimidinone nucleosides. Furthermore, the COSY spectra of HPLC Peak 1a (Figures 3.10 – 3.12) and HPLC Peak 2b (Figures 3.7 – 3.9) both possess the correlations characteristic of nucleosides.

The identification of the non-zebularine nucleoside products as furanosyl vs. pyranosyl sugars could be accomplished by analysis of their respective HMBC ^{1}H-^{13}C correlational NMR spectra, which provides data on the connectivity of the C1' sugar carbon to either the C4' or C5' carbon. However, in the absence of this data, the ^{1}H spectra provide several important clues regarding the identity of these 2-pyrimidinone nucleosides. In assigning the sugar configuration of HPLC Peak 1a (Figure 3.6B), the spectrum shows a large downfield shift in the C1' and C2' protons to approximately 6.15 and 4.55 ppm, respectively, which most closely resembles the chemical shifts for the α-

ribofuranosyl anomer (Kett, 1997; Tran-Dinh, 1977). However, the Δδ between the H5' and H5" resonances of approximately 0.2 ppm indicates that the nucleoside could be α-pyranosyl (Kett, 1997). Although the number of atoms in the sugar ring cannot be firmly established at this point, it appears that the base is attached to the sugar in the α configuration.

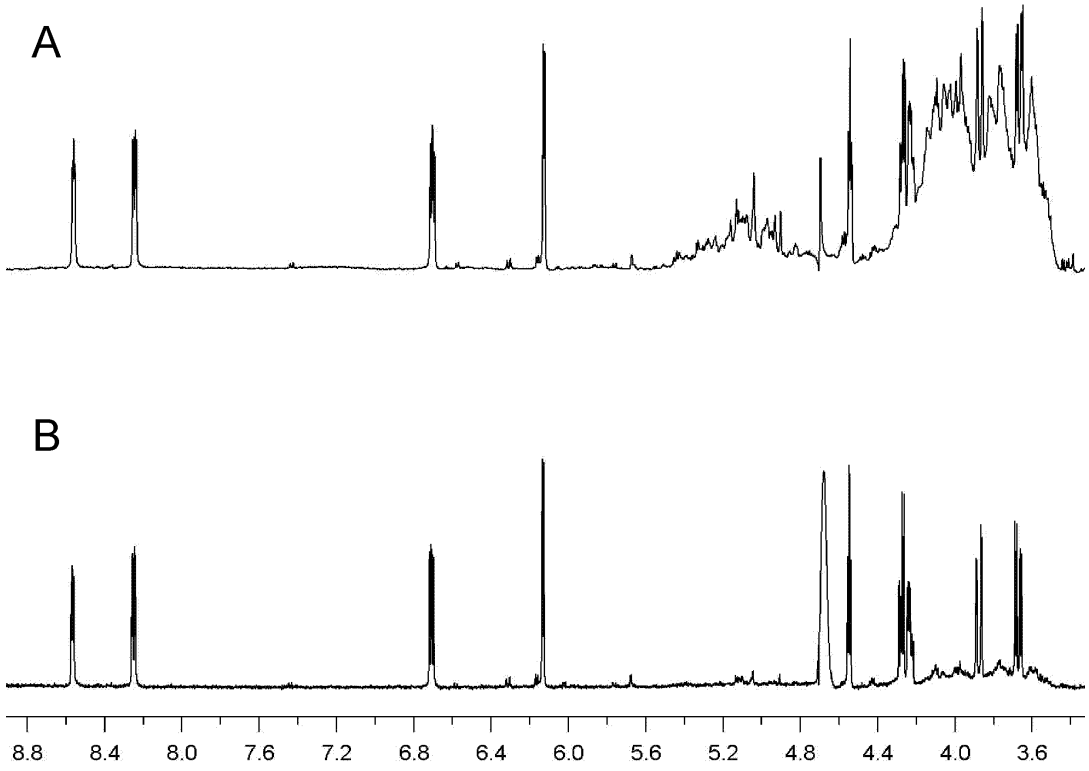

Figure 3.5 ^{1}H NMR spectra following the progress of HPLC purification of Peak 1a. (A) ^{1}H spectrum of HPLC Peak 1a after isolation under the conditions of an isocratic elution (Materials and Methods). The peaks in the region of 3.5 – 5.5 ppm resemble free (i.e. unreacted) ribose. (B) ^{1}H spectrum of HPLC Peak 1a after a second round of HPLC purification under the conditions of the gradient elution (Materials and Methods). This second step of purification reduced the concentration of the impurities in the sample enough to identify and assign the sugar protons of Peak 1a.

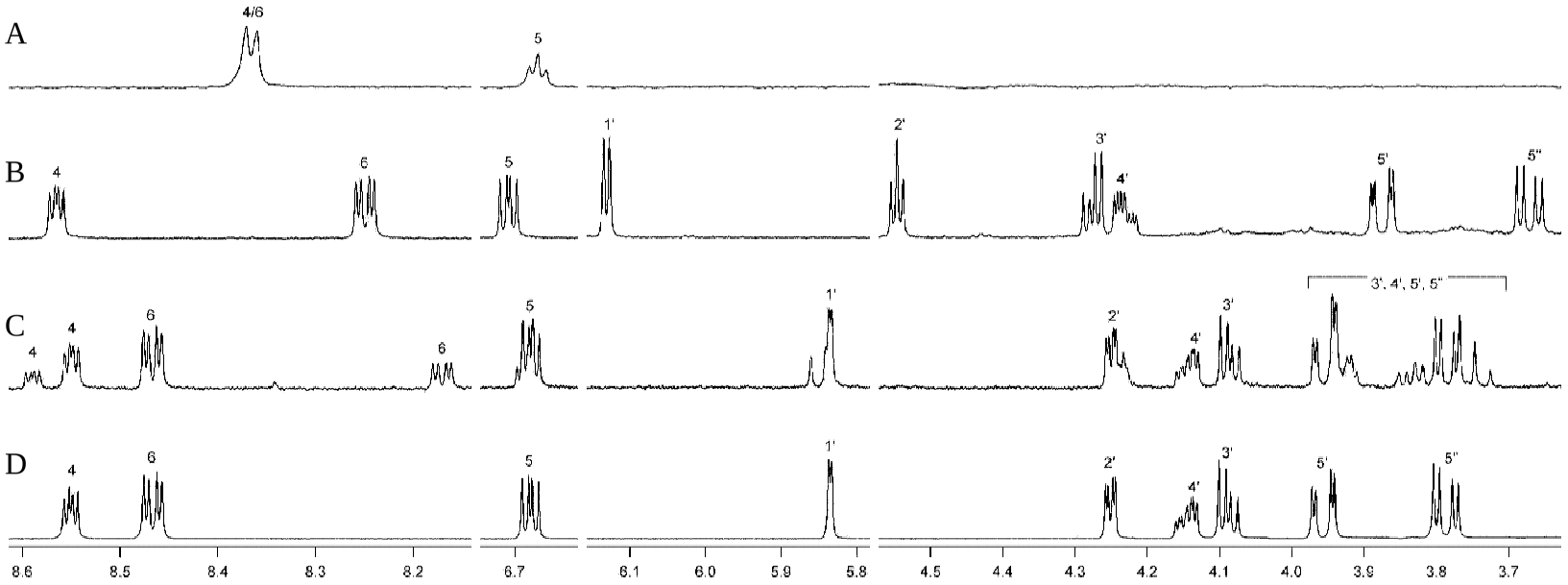

Figure 3.6. ^{1}H NMR spectra of the 2-pyrimidinone base and nucleoside standards and the HPLC purified nucleoside reaction products. Regions of the spectra that are not shown either contained no peaks or contained peaks from water or the HPLC solvent TEAA. (A) Spectrum of the 2-pyrimidinone base. The molecule is symmetric, giving rise to only two distinguishable proton peaks. (B) Spectrum of HPLC Peak 1a. (C) Spectrum of HPLC Peaks 2a and 2b, which were collected in the same HPLC fraction during purification. (D) Spectrum of the zebularine standard.

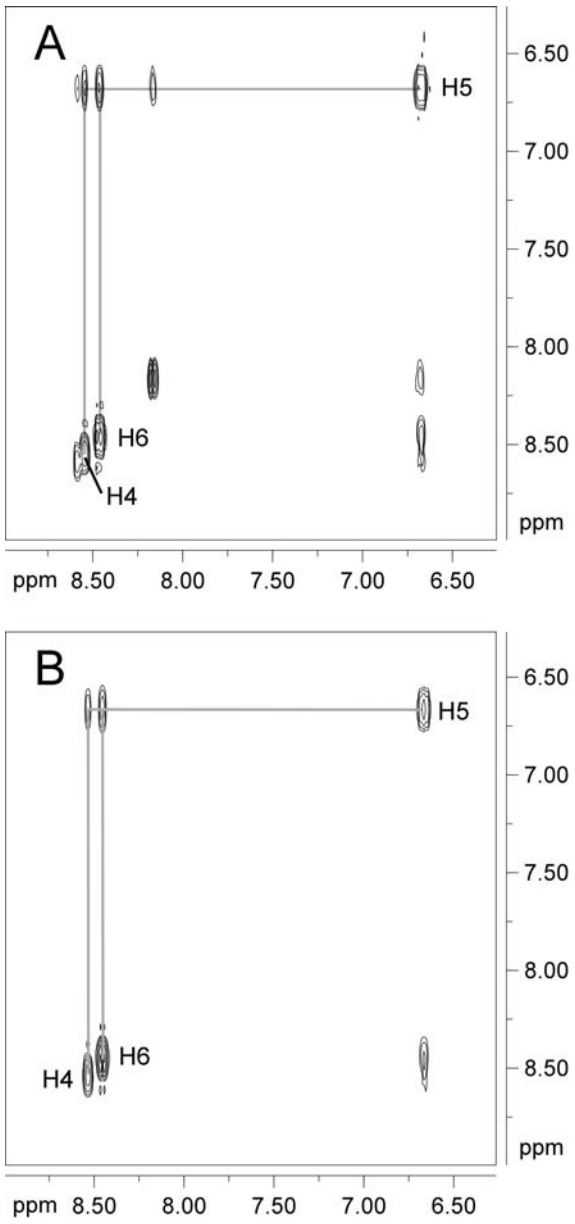

Figure 3.7 Comparison of COSY spectra for the base protons of HPLC Peak 2a and 2b (A) and the zebularine standard (B). Peak 2 contains two species, the more abundant of which is Peak 2a and has proton chemical shifts that match that of zebularine, as demonstrated by the correlation lines drawn. The less abundant species, HPLC Peak 2b, is also a 2-pyrimidinone nucleoside, with the H4 proton slightly downfield, H6 upfield, and H5 nearly coincident with their corresponding protons of zebularine.

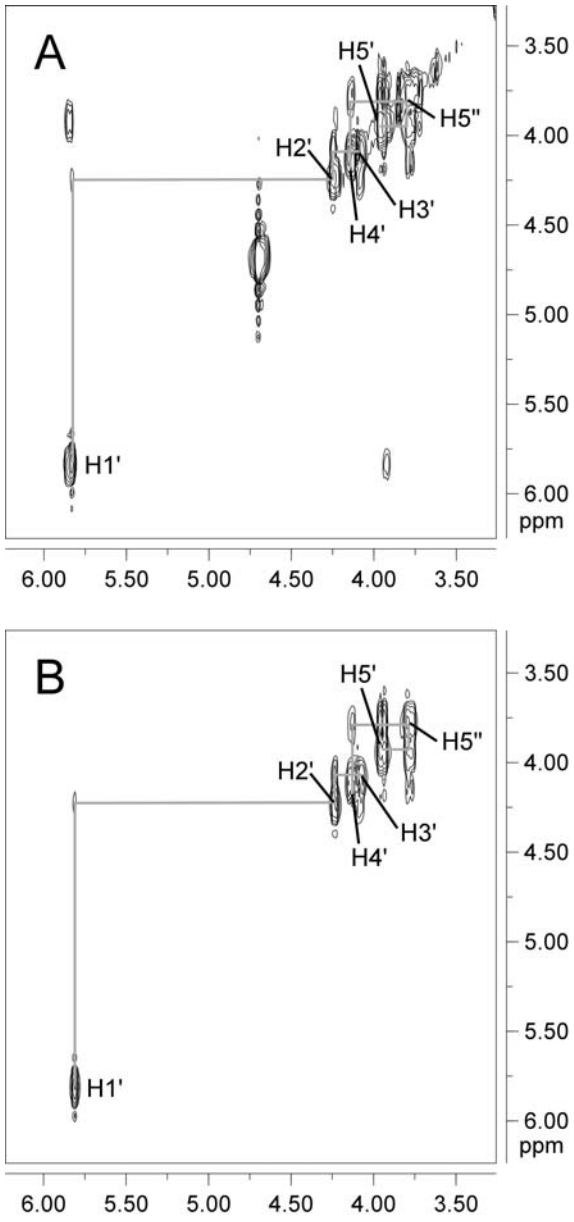

Figure 3.8 Comparison of COSY spectra for the sugar protons of HPLC Peaks 2a and 2b (A) and the zebularine standard (B). Peak 2 contains two species, the more abundant of which is Peak 2a and has proton chemical shifts that match that of zebularine, as demonstrated by the correlation lines drawn. The less abundant species, HPLC Peak 2b, is also a 2-pyrimidinone nucleoside, with the H1' nearly coincident with the H1' of zebularine.

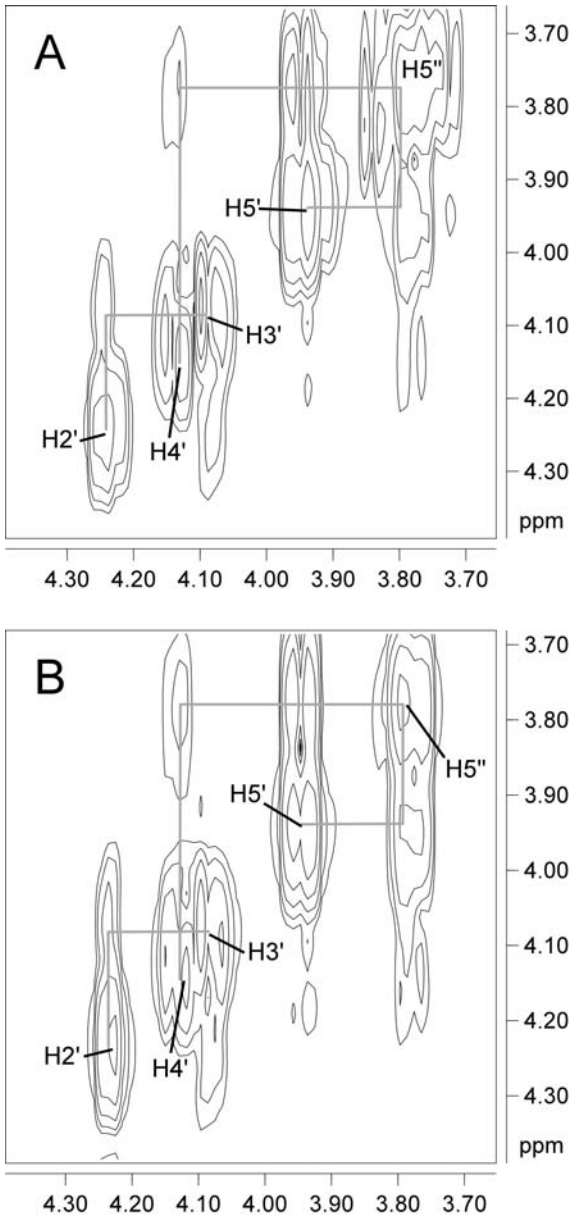

Figure 3.9 Comparison of COSY spectra for the region of the H2' – H5" sugar protons of HPLC Peaks 2a and 2b (A) and the zebularine standard (B). Peak 2 contains two species, the more abundant of which is Peak 2a and has proton chemical shifts that match that of zebularine, as demonstrated by the correlation lines drawn. The less abundant species, HPLC Peak 2b, is also a 2-pyrimidinone nucleoside, with the H2' proton nearly coincident with H2' of zebularine. H3' and H4' of Peak 2b are shifted upfield into the region with the H5' and H5" protons of both species.

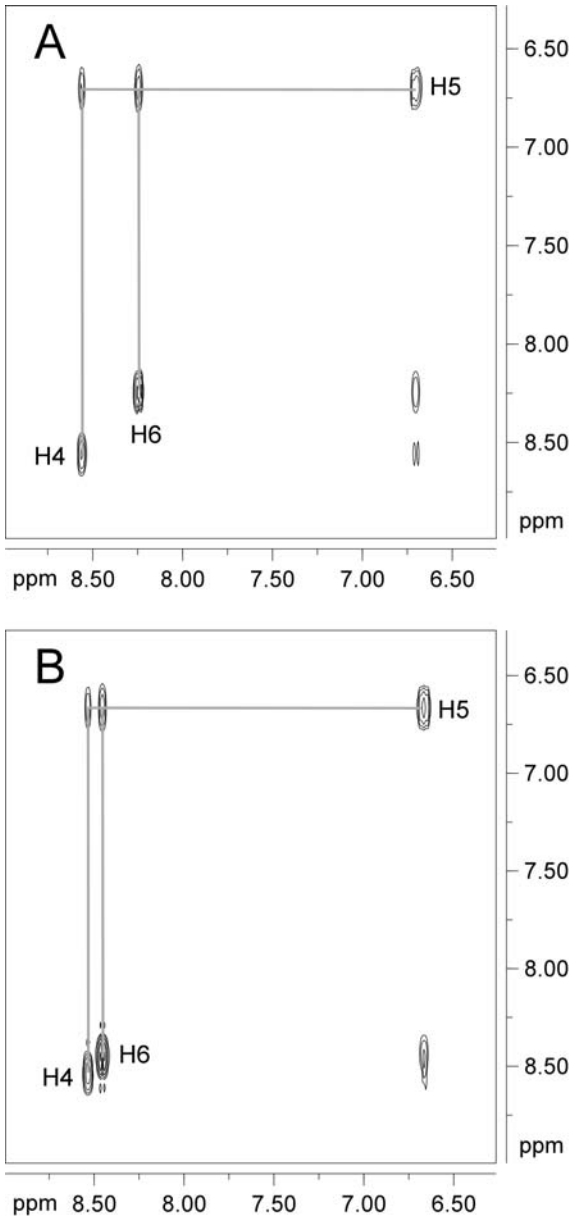

Figure 3.10 Comparison of COSY spectra for the base protons of HPLC Peak 1a, a 2-pyrimidinone nucleoside (A), and the zebularine standard (B). Note that the H4 and H5 protons of HPLC Peak 1a are at the same chemical shift as that for zebularine, whereas H6 is shifted upfield.

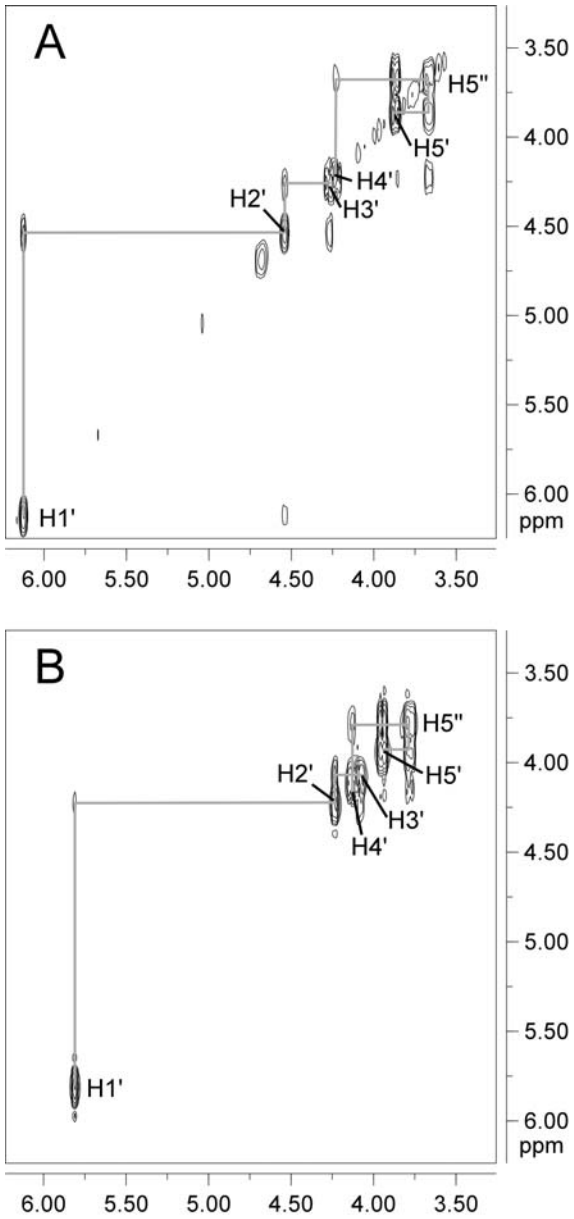

Figure 3.11 Comparison of COSY spectra for the sugar protons of HPLC Peak 1a, a 2-pyrimidinone nucleoside (A), and the zebularine standard (B). The H1' and H2' of HPLC Peak 1a are each shifted downfield from that of zebularine.

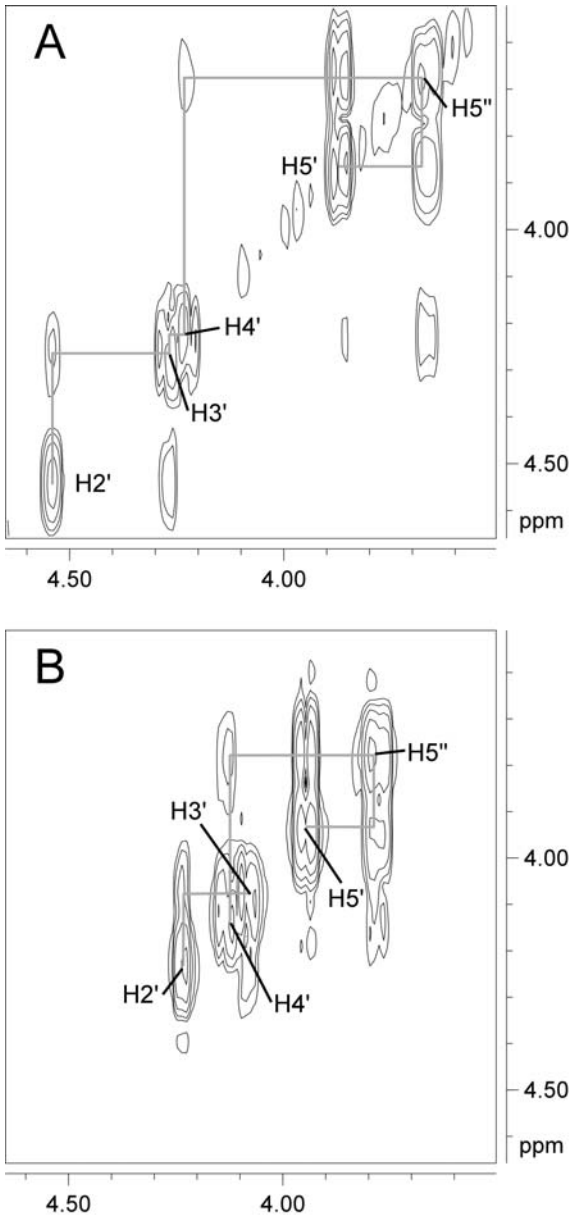

Figure 3.12 Comparison of COSY spectra for the region of the H2' – H5" sugar protons of HPLC Peak 1a, a 2-pyrimidinone nucleoside (A), and zebularine standard (B).

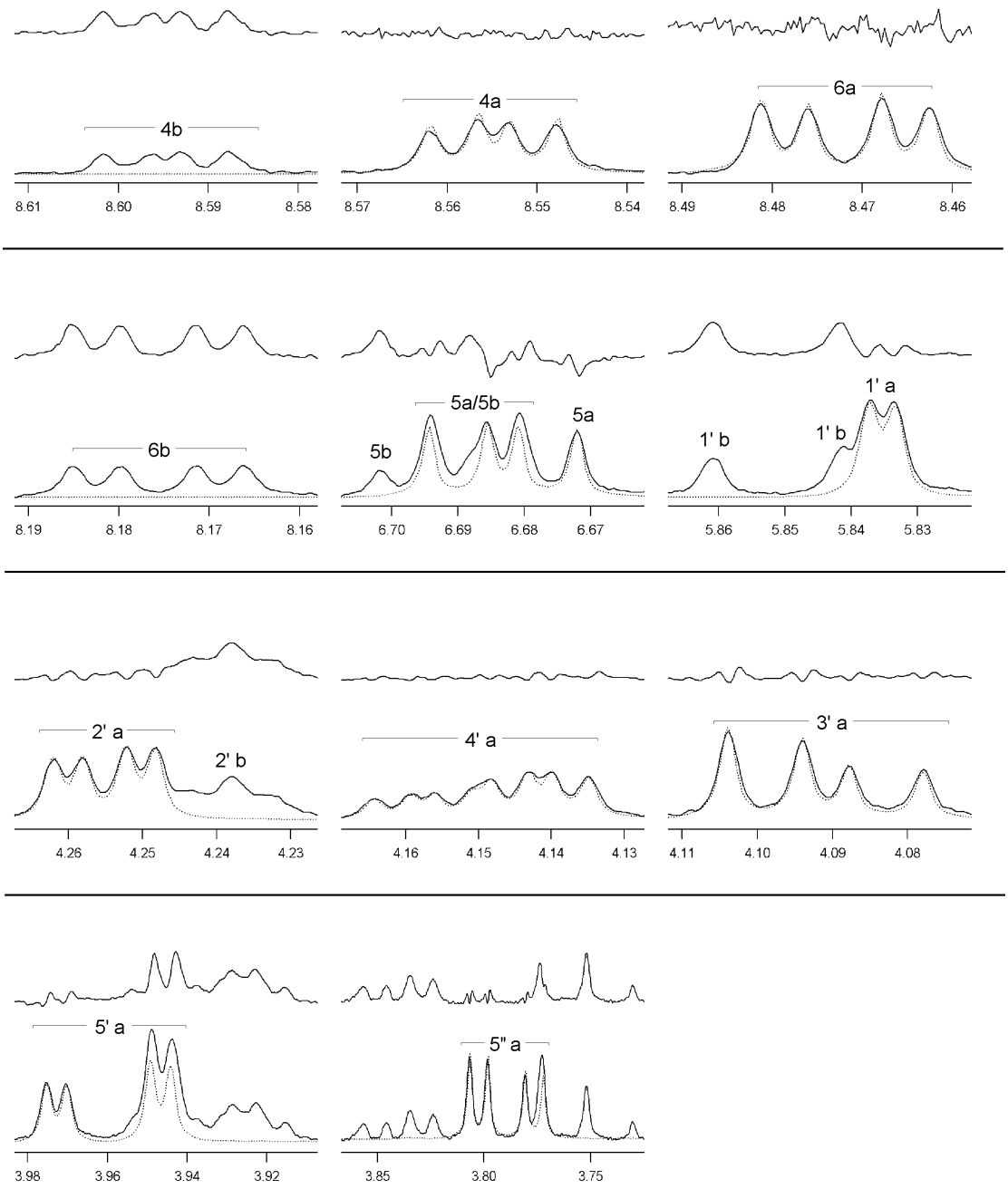

Figure 3.13 Comparison and subtraction spectra of HPLC Peaks 2a/2b with the zebularine standard. The bottom spectra of each set show the overlay of the zebularine standard (dotted line) with the nucleoside products. The zebularine proton resonances are a perfect match with the proton resonances of HPLC Peak 2a in both chemical shift and coupling constants. The top spectrum of each set shows the result of subtracting the zebularine NMR spectrum from the NMR spectrum of the nucleoside products, with the residual being the spectrum of HPLC Peak 2b. The resonance peaks labeled have been identified using the 2D COSY spectrum for each species.

Figure 3.13 shows the results of overlaying the NMR spectra of HPLC Peaks 2a/2b and the zebularine standard NMR spectrum (bottom of each set), as well as the residual spectrum that results from subtracting the zebularine spectrum from the HPLC product spectrum (top of each set). These comparisons provide additional information regarding these two nucleosides. First, the major product in the sample spectrum, HPLC Peak 2a, is irrefutably zebularine. Every chemical shift and every coupling constant of the authentic standard matches the more intense set of peaks in the sample, as well as published values for this compound (Driscoll, 1991). Second, the subtracted spectrum, which corresponds to HPLC Peak 2b, exhibits a coupling constant between sugar resonances H1' and H2' that is relatively large, approximately 10 Hz, and the H4' and H3' sugar resonances are shifted significantly upfield into the region where H5' and H5'' are typically found for ribosides (3.7 – 4 ppm). These data suggest that HPLC Peak 2b may be the β-ribopyranosyl form of the 2-pyrimidinone nucleoside (Kett, 1997; Wani, 1986).

3.4 Discussion

The successful prebiotic synthesis of zebularine, the β-ribofuranosyl nucleoside of 2-pyrimidinone, has been proven by HPLC (Peak 2a) and NMR analyses. In addition to zebularine, two other anomers of the nucleoside have been synthesized and isolated, which appear to be the β-ribopyranosyl nucleoside (HPLC Peak 2b) and one α-form ribonucleoside (HPLC Peak 1a). The HPLC data show that the distribution of the four expected nucleoside products does not resemble the equilibrium distribution of ribose anomers in solution (Figure 2.7). The reason for this difference is not immediately obvious. If the distribution of the nucleoside products depended on the equilibrium amounts of each conformation of free ribose, then the β-ribopyranosyl nucleoside would be the major product, rather than one of the two minor products. If steric hindrance were a factor, preventing the base from coming in close enough contact to the aldehydic

carbon for bond formation in certain sugar conformations, then it would be expected that the two β anomers would be the most abundant. However, of the two most abundantly synthesized species, one is β and one is apparently α. It has previously been established that the identity of the base in sugar-base condensation reactions can influence the conformation of the sugar in the resulting nucleoside (Fuller, 1972; Kett, 1997). Solution conditions of the condensation reaction can also affect the relative distributions of sugar conformations in the nucleoside products (Fuller, 1972). Taken together, this suggests that the ratios of products and the total amount of nucleosides formed in the prebiotic synthesis of nucleosides might be tunable, and could be optimized for producing the nucleoside conformer of interest.

4. OPTIMIZATION OF THE PYRIMIDINE NUCLEOSIDE-FORMATION REACTION

4.1 Introduction

In Chapter 3 it was demonstrated that 2-pyrimidinone forms nucleosides upon heating and drying with ribose. This result represents the first demonstration of spontaneous pyrimidine nucleoside formation by a simple drying reaction. HPLC and NMR analyses indicated that the four possible ribose nucleosides of 2-pyrimidinone (i.e. α-ribopyranosyl, α-ribofuranosyl, β-ribopyranosyl, and β-ribofuranosyl) are probably all formed in this reaction. Although only β-furanose nucleosides are used in present day RNA and DNA, other nucleoside anomers can form nucleic acid structures with stable base pairs (Eschenmoser, 1999). Thus, from a prebiotic chemistry perspective, all four nucleosides are of interest as potential ancestors to the nucleosides of contemporary life. In this chapter the effects of several changes in solution conditions on the 2-pyrimidinone nucleoside-formation reaction are explored. Of particular interest are reaction conditions that improve nucleoside yield, and conditions that lead to the enhancement of one specific nucleoside anomer over the other three. Additionally, observing how changes in reaction conditions affect nucleoside yield could provide valuable insights into the mechanism of the reaction, and provide clues to how pyrimidine nucleosides could have first appeared on the prebiotic Earth.

While the results described in the previous chapter are very promising, it is not expected that life emerged from a solution with a pH of 2.5 (Bengston, 1994), which is the pKa of the 2-pyrimidinone and the pH of an unbuffered sample solution. Thus, it is important to explore the ability for 2-pyrimidinone to form nucleosides with ribose over a broader range of pH conditions. It is generally accepted that the purine nucleoside formation reaction described by Orgel and co-workers (Fuller, 1972) is acid catalyzed

and proceeds by the nucleophilic attack of the purine base on the C1' of ribose (Ingar, 2003). In solutions at neutral pH, first shell water molecules of metal ion hydrates can act as Lewis acids and donate protons in acid-catalyzed reactions. Metal ions also have the potential to directly chelate the –OH groups of sugars. Accordingly, previous investigations of purine nucleoside synthesis under plausible prebiotic conditions have suggested that metal ion salts played an important role as prebiotic catalysts (Fuller, 1972). It is possible that cations of different valence and size would preferentially chelate different anomers of ribose and thereby preferentially catalyze the synthesis of nucleosides with a particular anomeric form of ribose. Thus, testing the effect of different salts and their concentrations on the 2-pyrimidinone nucleoside-formation reaction could be crucial to understanding this reaction and its potential to be considered plausibly prebiotic.

Dehydration reactions appear to be the main mechanism by which small molecules on the prebiotic Earth became part of the original polymers of life (Schopf, 1999). Multimolecular reactions of soluble substrates can, of course, be dramatically enhanced during water evaporation as substrates that remain in solution increase in concentration until solubility limits are reached. In the case of the 2-pyrimidinone nucleoside-formation reaction, the significance of dehydration in driving this reaction forward can be tested by carrying out experiments in sealed vessels that prevent drying of the reaction mixtures.

Reaction temperature is undoubtedly an important factor for optimization of this nucleoside-formation reaction. Analyzing product yields for reactions carried out at a range of temperatures can reveal the point at which nucleoside production decreases due to ribose or base degradation. The loss of ribose at elevated temperatures is expected to be a particular limitation in nucleoside-formation reactions, as Miller and co-workers have shown that the half-life of ribose can decrease to a matter of minutes at 100ºC, pH 7.0 (Larralde, 1995). Experiments carried out with different ribose concentrations could

be used to determine if ribose degradation is a limiting factor in the 2-pyrimidinone nucleoside formation reaction.

Overall, it is anticipated that understanding the role played by external conditions such as temperature, dehydration, and surface area and internal solution conditions such as reaction pH, salt composition, and salt concentration can lead to a marked improvement in the efficiency of the pyrimidine nucleoside reaction. Optimized and plausibly prebiotic conditions for these reactions could prove to be valuable tools for the quest to solve "The Nucleoside Problem".

4.2 Materials and Methods

4.2.1 Sample preparation and HPLC analysis

Unless otherwise noted, aqueous samples of 5 mM 2-pyrimidinone and 75 mM ribose were prepared as 100 μl reaction volumes and the pH was measured and adjusted, if necessary. Salt composition and concentration, reaction temperature and the reaction vessel were varied in the following experiments, as detailed below. Reaction products were analyzed by HPLC on a Phenomenex ODS analytical column (4.6 mm x 250 mm, 5 μ) with mobile phase conditions of 10-30% B (0-5 min), 30-50% B (5-30 min), flow rate = 1.0 ml/min. Mobile phase A was 0.1 M triethyl ammonium acetate (TEAA), and mobile phase B was 25% (v/v) acetonitrile, 75% (v/v) 0.1 M TEAA. Elution products were monitored by UV absorbance at 303 nm as a function of retention time from 0 to 36 minutes.

4.3 Results

4.3.1 Equilibrium Experiments

The importance of dehydration in facilitating nucleoside formation was investigated by conducting experiments in which reaction mixtures were kept in sealed

vials to prevent loss of water. Each reaction contained 2.88 M 2-pyrimidinone base and 2.88 M ribose. The sealed reaction vials were placed at several temperatures (4, 22, 37, 65 or 85˚C). After 22 hr, the samples were diluted with dH_2O, and products were analyzed by HPLC according to the standard gradient and run time.

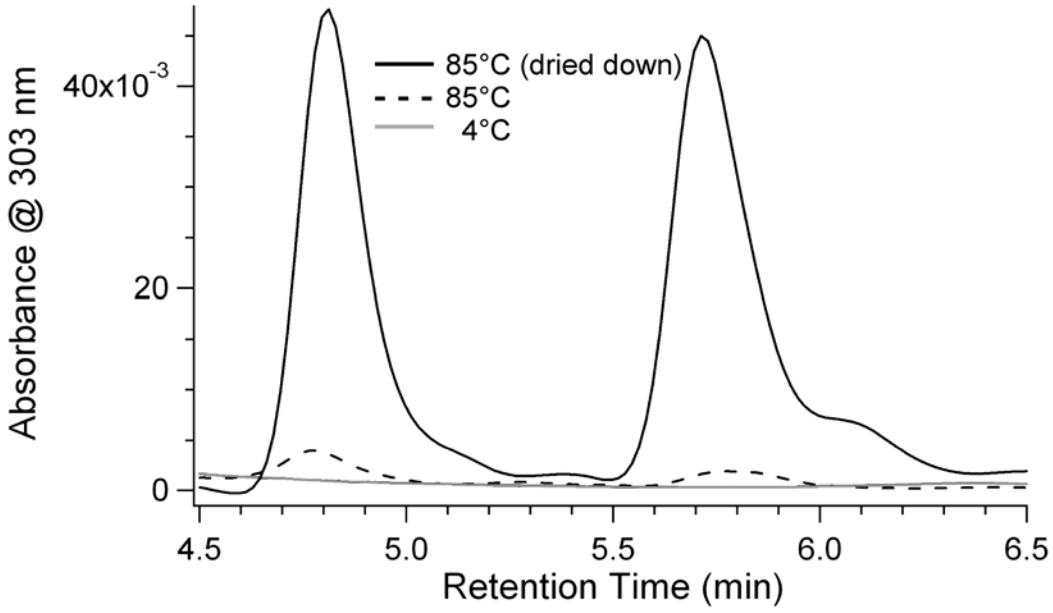

Figure 4.1. HPLC chromatographs illustrating the effect of sample drying on the nucleoside-formation reaction. Reaction components: 2.88 M 2-pyrimidinone base and 2.88 M ribose. The reactions with chromatographs labeled 4˚C (gray line) and 85˚C (dashed line) were sealed in vials for 22 hr. The reaction corresponding to the chromatograph labeled 85˚C (dried down) contained 5 mM 2-pyrimidinone base and 75 mM ribose and was dried at 85˚C on a watch glass (black solid line).

Figure 4.1 shows the results of nucleoside yield (HPLC peaks at ~4.8 and 5.8 min) as a function of reaction temperature. Only two results, 4 and 85˚C, are shown for clarity. Samples maintained at temperatures 65˚C or less without drying (i.e. 4, 22, 37, 65˚C) showed no products or substantially less products than the sample maintained at 85˚C. The chromatograph for the 4˚C sample is shown in Figure 4.1 for comparison. The data indicate that the general trend is that increased temperatures yield more nucleoside

product in reactions that are sealed to prevent dehydration (Figure 4.1). However, the yield of nucleoside is also affected by the extent of drying during the incubation. For comparison, the chromatograph of a reaction mixture dried at 85˚C and with only a fraction of the amount of the starting materials, (5 mM 2-pyrimidinone base and 75 mM ribose) is also shown in Figure 4.1. The formation of nucleoside products of the 85˚C reaction that was not dried down indicates that 2-pyrimidinone can form nucleosides with ribose under equilibrium conditions. However, a qualitative comparison of reactions carried out with and without drying indicates that drying substantially increases nucleoside yield.

4.3.2 Effect of Temperature on Dehydration Reactions

Since it was established in the previous section that both temperature and dehydration influence nucleoside yield, the effect of different drying temperatures on reaction product formation was investigated. Each aqueous reaction contained 5 mM 2-pyrimidinone base, 75 mM ribose, 25 mM $MgCl_2$, and 50 mM $MgSO_4$. Reactions were prepared on a watch glass, placed in an oven to dry (65 or 85˚C, 2 hr), rehydrated in dH_2O, and the products were analyzed by HPLC according to standard gradient conditions and run time.

Nucleoside products, corresponding to HPLC peaks at 4.9 and 5.9 min column retention times in Figure 4.2, are enhanced at 85˚C in comparison to 65˚C. These results suggest that the extent of dehydration at 65˚C is not sufficient to give maximum nucleoside product, or that higher temperatures are necessary to overcome a kinetic barrier. Even though the rate of glycosidic bond formation is enhanced at 85˚C, increased degradation of ribose sugar may also be taking place at elevated temperature conditions.

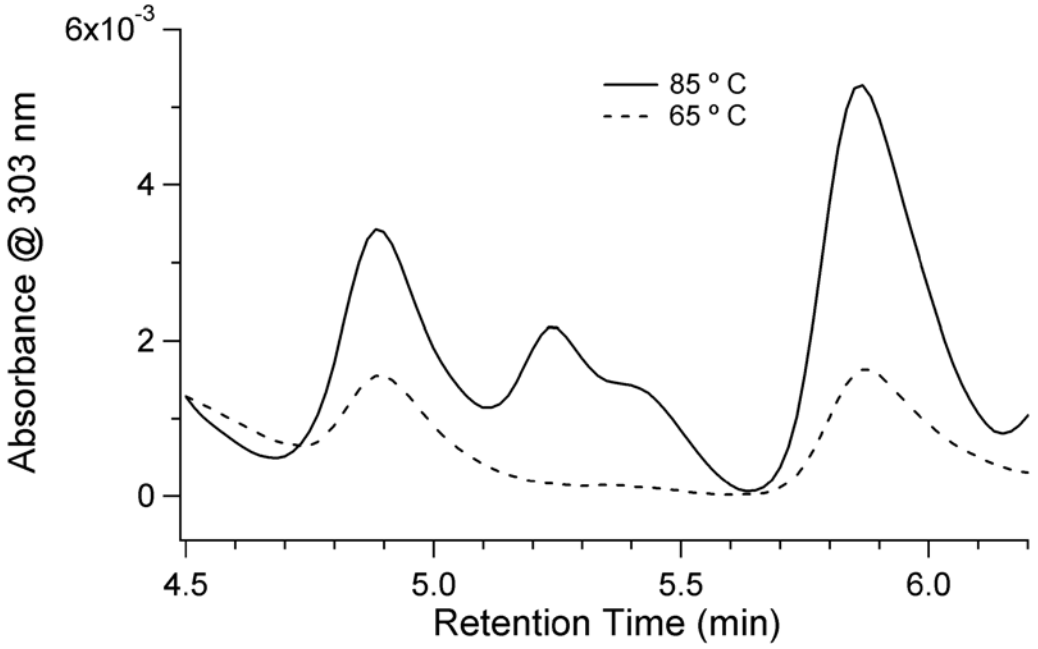

Figure 4.2. HPLC chromatographs illustrating the effect of temperature on the nucleoside reaction. Reaction components: 5 mM 2-pyrimidinone base, 75 mM ribose, 25 mM $MgCl_2$, and 50 mM $MgSO_4$. Reactions were prepared on a watch glass and dried at 65˚C (dashed line) or 85˚C (solid line) for 2 hr.

4.3.3 Effect of Surface Area on Dehydration

The effect of the vessel used in the drying process (watch glass or microcentrifuge tube) on nucleoside product formation was studied. Each reaction contained 5 mM 2-pyrimidinone base and 75 mM ribose. 100 µl reaction mixtures were prepared and either placed on a watch glass or in an open 1.5 ml Eppendorf microcentrifuge tube and placed in oven to dry at 85˚C for 2 hr. The samples were subsequently rehydrated in dH_2O and the products were separated by HPLC according to the standard gradient and run time.

The rate of sample dehydration is expected to be roughly proportional to the surface area of the reaction container. It is therefore expected that reaction solutions placed on watch glasses would experience increased dehydration (i.e. great water loss) or

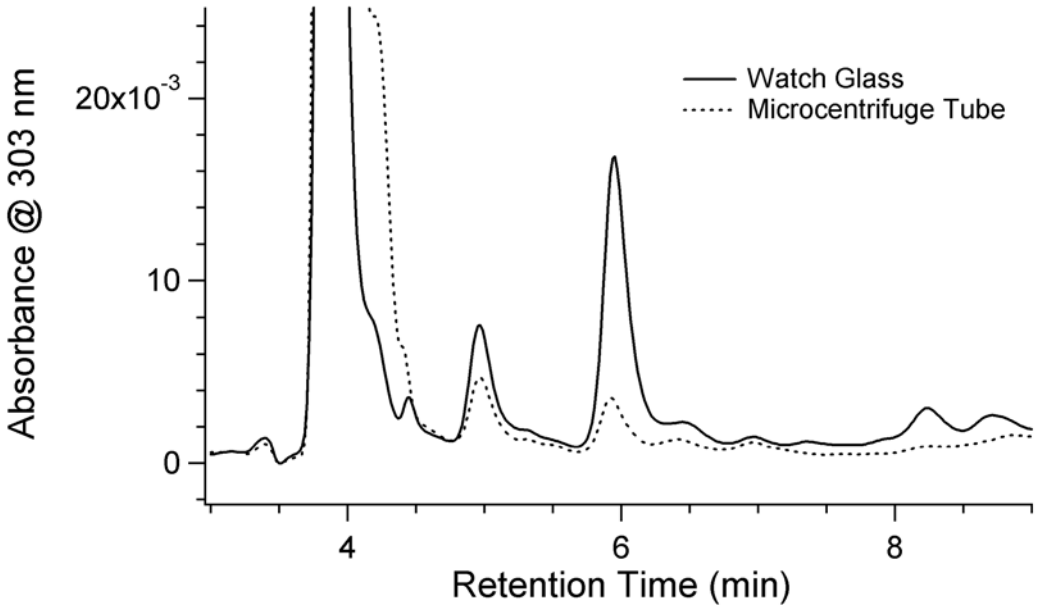

Figure 4.3. HPLC chromatographs illustrating the effect of surface area on the nucleoside-formation reaction. Reaction mixtures contained 5 mM 2-pyrimidinone base and 75 mM ribose. Reactions were prepared either on a watch glass (solid line) or in 1.5 mL microcentrifuge tube (dashed line).

longer times at the terminal level of dehydration in comparison to samples in microcentrifuge tubes. Products peaks at ~5 and 6 min column retention times in Figure 4.3 correspond to the nucleoside products and indicate that significantly higher product formation is observed when reactions were conducted on a watch glass. The area measured under the HPLC peak at ~4 min is also consistent with a greater quantity of intact (not degraded) and unreacted 2-pyrimidinone base remaining when the reaction solution was placed in microcentrifuge tubes for drying.

4.3.4 Modulating the Reaction pH

The effect of solution pH on the 2-pyrimidinone nucleoside-formation reaction was also tested. Each reaction contained 5 mM 2-pyrimidinone base and 75 mM ribose. Reactions were adjusted to pH 7 by the addition of base, whereas the reaction solution at pH 2.5 corresponds to the reaction with unadjusted pH. Reactions were prepared on a watch glass and placed in an oven to dry at 100˚C for 2.25 hr. Reaction mixtures were

rehydrated in dH_2O and the products were separated by HPLC according to standard gradient and run time.

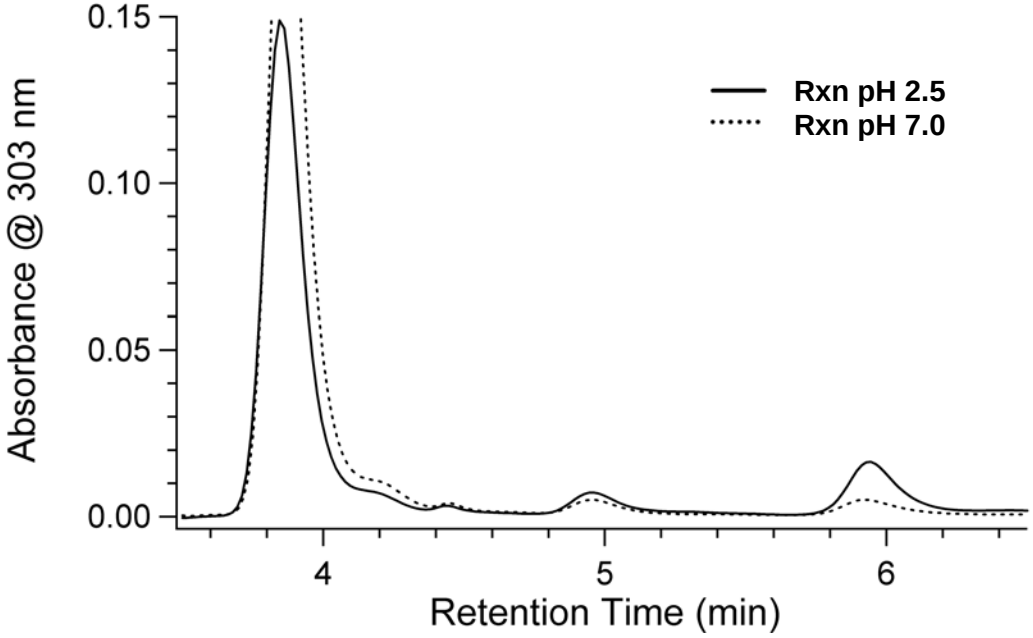

Figure 4.4. HPLC chromatographs illustrating the effect of pH on the nucleoside-formation reaction. Reaction mixtures contained 5 mM 2-pyrimidinone base and 75 mM ribose. Reactions at pH 2.5 (solid) or pH 7.0 (dotted).

Nucleoside HPLC peaks at ~5 min and ~6 min column retention times in Figure 4.4 correspond to the nucleoside products and show that maximum product formation occurs at an unadjusted pH of 2.5. The reaction at pH 7.0 (dotted line in Figure 4.4) also shows a significant amount of unreacted base (peak at ~4 min) which suggests that the decline in product formation cannot be attributed to base degradation. These results suggest that either the reaction is acid catalyzed or that the pKa of 2-pyrimidinone base (i.e. 2.23) is an important parameter to consider when altering the solution pH of these reactions.

4.3.5 Dehydration Reactions in the Presence of Various Salts

The effect of varying the cation identity in the chloride salt of the reaction was also studied. Eight different reaction mixtures were prepared such that the final

concentrations of substrates were 5 mM 2-pyrimidinone base, 75 mM ribose, and 75 mM metal chloride salt. The metal ions tested were Li^{+}, Na^{+}, Mg^{2+}, Ca^{2+}, Sr^{2+}, Fe^{2+}, Ni^{2+} and Zn^{2+}. Reactions were prepared on a watch glass, placed in oven at 85˚C for 2 hr, rehydrated in dH_2O, and the products were analyzed by HPLC using the standard HPLC gradient conditions described above.

Table 4.1. The effect of metal salts on the nucleoside-formation reaction. Reaction components: 5 mM 2-pyrimidinone base, 75 mM ribose, 75 mM chloride salt. Yields provided are percentages, normalized to the yield of the salt-free (pH 2.5) β-ribofuranosyl product (HPLC column retention time 5.8 min). Samples were dried at 85˚C for 2 hr.

Salt	HPLC peak at 5.8 min	HPLC peak at 6.4 min
No added salt (pH 2.5)	100	7
$MgCl_2$	38	5
LiCl	23	2.7
NaCl	8	< 1
$CaCl_2$	7	< 1
$FeCl_2$	6	< 1
$SrCl_2$	5	< 1
$NiCl_2$	6	< 1
$ZnCl_2$	5	< 1

The results presented in Table 4.1 demonstrate that, of the metal chlorides tested in this study, Mg^{2+} is the most effective cation for catalyzing the glycosidic bond formation between the ribose sugar and the 2-pyrimidinone base. It is expected that the pH, or more strictly the hydronium ion activity, during the final stages of sample drying will be governed by the concentration and species of salt used in these reactions. The data shown here suggest that the formation of nucleoside products can be modulated by the species of metal ion present in the reaction. The observed effects are presumably indicative of the ability of the metal ion hydrates to act as Lewis acid catalysts and/or the metal ions' abilities to coordinate hydroxyl groups on the sugar, affecting the sugar configuration.

4.3.6 Effect of $MgCl_2/MgSO_4$ Mixed Salt Concentrations on the Nucleoside-Formation Reaction

Orgel and co-workers have demonstrated that the most efficient synthesis of purine nucleosides occurred in a natural seawater environment, which contains $MgCl_2$ and $MgSO_4$ in a 1:2 ratio (Fuller, 1972). The work presented above also indicates that Mg^{2+} is the most effective metal ion for the synthesis of pyrimidine nucleosides with 2-pyrimidinone. Together, these results suggested that it would be of interest to investigate the effect of the absolute concentration of a 1:2 mixture of $MgCl_2$ and $MgSO_4$ on the 2-pyrimidinone-ribose reaction. In this study, each reaction mixture contained 5 mM 2-pyrimidinone base, 75 mM ribose, and absolute concentrations of $MgCl_2$, $MgSO_4$ of 0, 0; 5, 10; 10, 20; 25, 50; 50, 100; 100, 200; and 250, 500 mM, respectively. Reactions were prepared on a watch glass, placed in an oven to dry (85˚C, 2 hr), rehydrated in dH_2O, and the products were analyzed by HPLC using the standard gradient conditions described above.

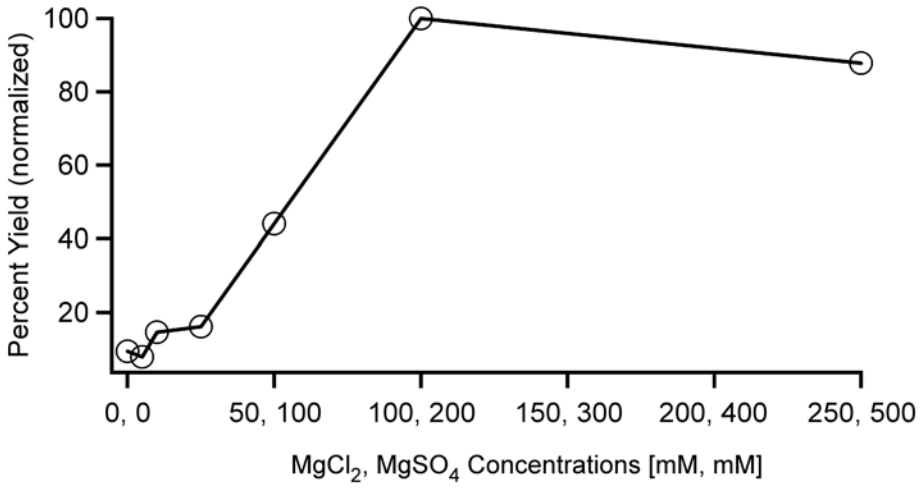

Figure 4.5. Yield of the β-furanose nucleoside as a function of $MgCl_2$, $MgSO_4$ salt concentrations. Yields were normalized to the reaction with 100, 200 mM $MgCl_2$, $MgSO_4$. Reaction components: 5 mM 2-pyrimidinone base, 75 mM ribose and varying total concentrations of $MgCl_2$, $MgSO_4$ (0, 0; 5, 10; 10, 20; 25, 50; 50, 100; 100, 200; and 250, 500 mM).

An increase in β-furanose nucleoside yield was observed when the $MgCl_2$, $MgSO_4$ salt mixture concentration was increased from 0, 0 mM to 100, 200 mM (Figure 4.5). It is possible that these particular magnesium salts act both as Lewis acids and aid in the dehydration process to facilitate nucleoside formation.

4.4 Discussion

Based on the analyses presented above for various experimental conditions, formation of the β-ribofuranosyl nucleoside product by the reaction of 2-pyrimidinone base with ribose is optimal for samples containing 5 mM 2-pyrimidinone base and 75 mM ribose under the reaction conditions of:

- 100 mM $MgCl_2$, 200 mM $MgSO_4$ (before dehydration)
- Dehydrating conditions (85˚C on a watch glass)
- pH ~2.5

The experimental conditions outlined above are likely to be somewhat specific for the nucleoside reaction of 2-pyrimidinone base with ribose. Although further analyses would need to be performed to determine exact yields provided by these reaction conditions, the UV data gathered after HPLC purification suggest that nucleoside yield is greater than 1%. It remains to be demonstrated whether these conditions can be generally applied to nucleoside reactions with other pyrimidine bases and other sugars. Some of these experiments are discussed in the next chapter of this thesis.

5. INITIAL INVESTIGATIONS ON THE MECHANISM AND GENERALITY OF THE PYRIMIDINE NUCLEOSIDE FORMATION REACTION

5.1 What is the Optimum pH for 2-Pyrimidinone Nucleoside Formation?

Orgel and co-workers have previously demonstrated that glycosidic bond formation is acid catalyzed when the purine bases are heated and dried in the presence of ribose (Fuller, 1972). These investigators added either HCl or various salts to solutions of purine bases and ribose that were then heated and dried. HCl was found to be one of the best catalysts tested (Fuller, 1972). In experiments carried out at neutral pH, the hydrates of divalent metal ions apparently acted as general acid catalysts. It is well known that DNA and RNA depurination is acid catalyzed. Thus, it was perhaps not unexpected that purine nucleoside synthesis could take place at near-anhydrous conditions by the reversal of the acid-catalyzed depurination reaction.

Most of the nucleoside synthesis reactions involving 2-pyrimidinone described in this thesis were carried out at mildly acidic pH (i.e. around pH 2.5). The reason for this choice of pH was originally the result of a very practical consideration. Commercially available 2-pyrimidinone is sold as a hydrochloride salt. 2-pyrimidinone has a pKa of 2.24 (Albert, 1956). Reaction solutions prepared with 2-pyrimidinone and ribose, without additional buffers, had a measured pH of approximately 2.5. At this pH, approximately one half of the 2-pyrimidinone molecules would have both nitrogens protonated.

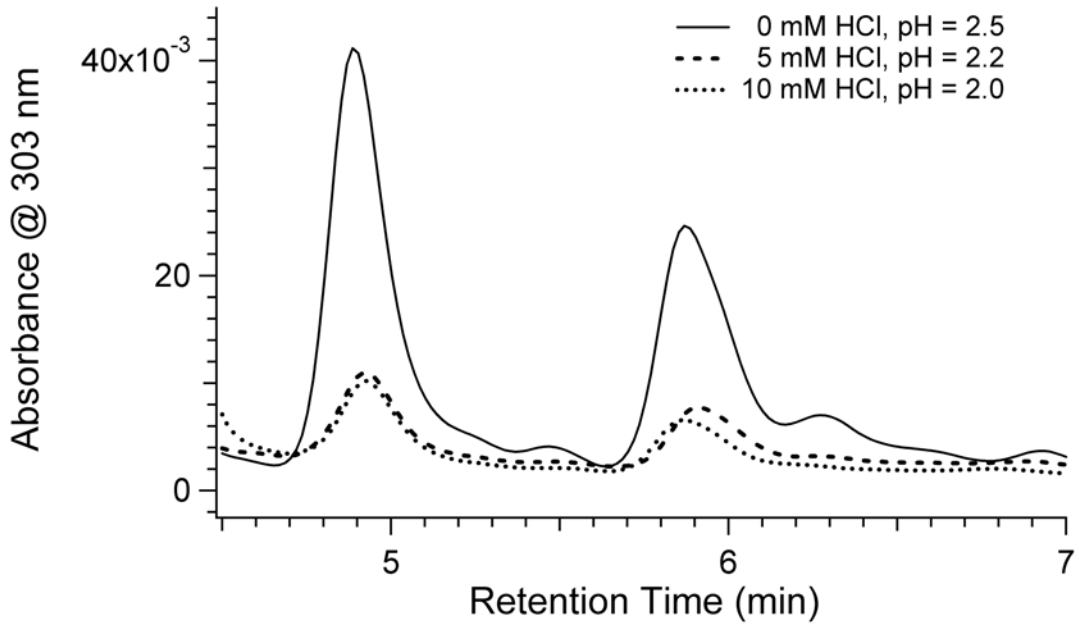

Figure 5.1: HPLC chromatographs illustrating the effect of pH on nucleoside formation by 2-pyrimidinone and ribose. Reaction mixtures contained 5 mM 2-pyrimidinone base and 75 mM ribose. Reactions were prepared on a watch glass and placed in an oven at 85˚C for 2 hr. Reaction mixtures were rehydrated in dH_2O and the products were analyzed according to standard HPLC gradient conditions. Reactions were at pH 2.5 (solid, 0 mM HCl), pH 2.2 (dashed, 5 mM HCl), and pH 2.0 (dotted, 10 mM HCl).

To test if the 2-pyrimidinone nucleoside-formation reaction is acid catalyzed, HCl was added to reaction mixtures to final concentrations of 10 mM and 20 mM. These additions of HCl reduced the pH of the reaction solutions from pH 2.5 to pH 2.2 and 2.0, respectively. Samples were then dried, along with a control sample at pH 2.5, in an oven on watch glasses at 85ºC. It was found that reducing the pH of the reaction solution by even 0.3 pH units resulted in reduced nucleoside yield (Figure 5.1). Thus, the data presented here indicates that the formation of zebularine from 2-pyrimidinone is very sensitive to pH. Iocono et al. have reported that the glycosidic bond of the 2-pyrimidinone-2'-deoxynucleoside (i.e. 2'-deoxy-zebularine) incorporated into a DNA strand is cleaved within 24 hr at pH 3.0 and ambient temperature (Iocono, 1990). This result suggests that 2-pyrimidinone coupling to a sugar could be acid catalyzed around

pH 3.0. However, as reported in previous chapters of this thesis, zebularine formation was not observed unless samples were heated during drying. Thus, it appears that either cleavage of the glycosidic bond of 2'-deoxy-zebularine proceeds by a different mechanism than the formation of zebularine, or zebularine is more kinetically or thermodynamically stable than 2'-deoxy-zebularine. In either case, it appears that the pyrimidine nucleoside-formation reaction reported in this thesis proceeds by a reaction mechanism that is distinct from that described earlier by Orgel and co-workers for the synthesis of the purine nucleosides (Fuller, 1972).

5.2 Can Bases Similar to 2-Pyrimidinone Form Nucleosides by the Same Reaction?

The coupling of 2-pyrimidinone to ribose, as described in this thesis, represents the first successful production of a pyrimidine nucleoside in a plausible prebiotic reaction. Thus, it is of great interest to determine what chemical features of 2-pyrimidinone are essential for this reaction to occur. As a fist step toward answering this question, pyrimidine molecules related to 2-pyrimidinone were also tested for their ability to form nucleosides with ribose under similar reaction conditions. The structures of the pyrimidines tested thus far are shown in Figure 5.2

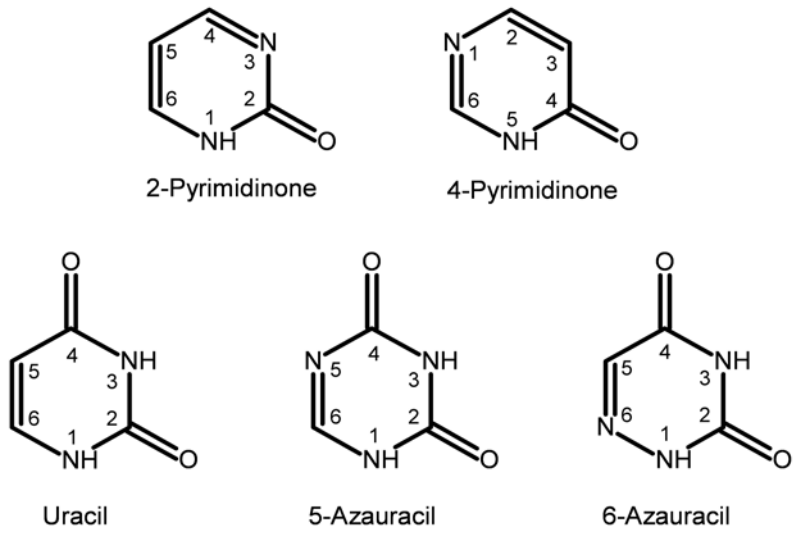

Figure 5.2: Pyrimidines tested in the nucleoside-formation reaction with ribose.

The pyrimidine most similar to 2-pyrimidinone is 4-pyrimidinone (Figure 5.2). This base has exactly the same chemical composition as 2-pyrimidinone, but with the keto group being adjacent to one nitrogen atom rather than being in between the two nitrogen atoms of the pyrimidine ring. 4-pyrimidinone has a pKa of 1.84, compared to the 2.24 pKa of 2-pyrimidinone (Albert, 1956). Unlike 2-pyrimidinone, 4-pyrimidinone is commercially available as a neutral reagent. An aqueous solution prepared with 4-pyrimidinone and ribose for the nucleoside reaction had an initial pH of 5.0. Unlike 2-pyrimidinone, 4-pyrimidinone is not symmetric about the keto group, which implies that there are two distinct positions on the base at which a glycosidic bond could be formed (i.e. N1 and N3). Thus, there are theoretically eight possible nucleosides that could be formed between 4-pyrimidinone and ribose. Although authentic standards are not commercially available for any of the 4-pyrimidinone nucleosides, it is expected that the nucleosides with 4-pyrimidinone connected to ribose through N3 would have HPLC retention times similar to those of the 2-pyrimidinone nucleosides.

In Figure 5.3 HPLC chromatographs are shown for the products formed by 4-pyrimidinone upon heating and drying with ribose at 85˚C for 2 hr. The reaction carried out at pH 5.0 exhibited HPLC peaks with retention times that were similar to those of HPLC peak 1 and peak 2 described earlier for 2-pyrimidinone. However, in comparison to peak 1 and peak 2 of the 2-pyrimidinone/ribose reaction carried out at pH 2.5, the yield of the 4-pyrimidinone reaction at pH 5.0 was approximately 40-fold less. When the pH of the 4-pyrimidinone reaction was adjusted to pH 2.4 by the addition of HCl (to have a pH comparable to that of the 2-pyrimidinone reaction), the yield of the putative 4-pyrimidinone nucleosides increased to approximately one-fourth of the yield of the 2-pyrimidinone reaction (Figure 5.3). A further decrease in pH to 2.1 resulted in a decrease in product yield. Thus, it appears that 4-pyrimidinone is also able to form nucleosides with ribose, and that the reaction is most favored under conditions very similar to those identified for 2-pyrimidinone.

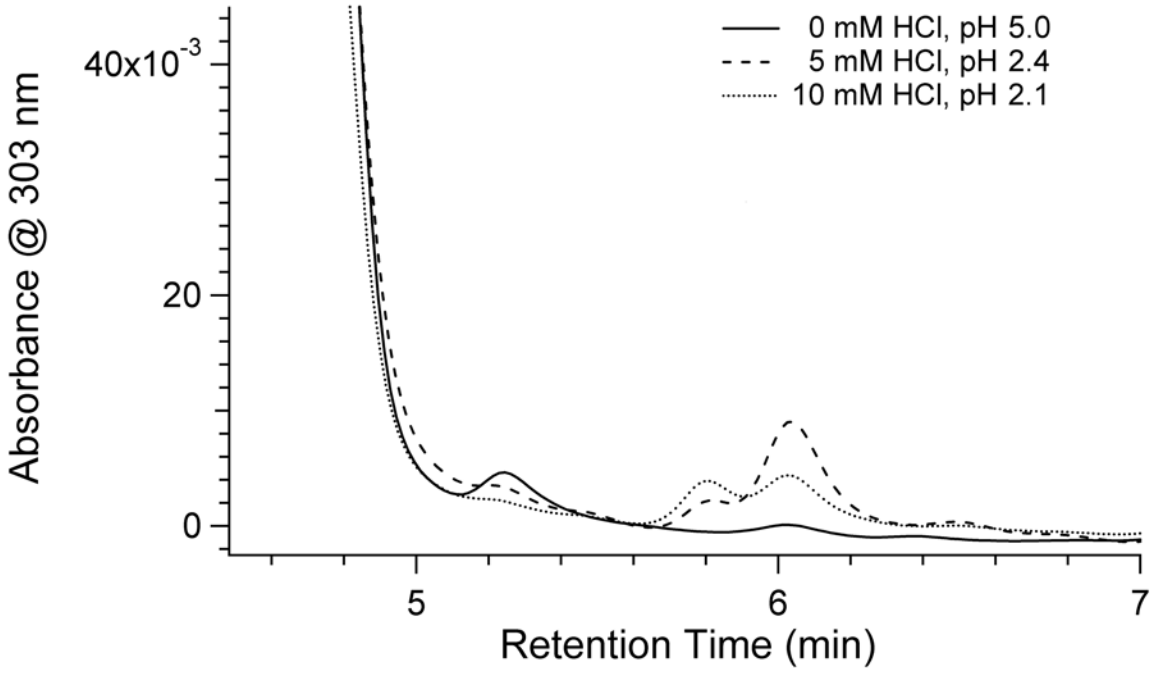

Figure 5.3: HPLC chromatographs illustrating the effect of pH on nucleoside formation by reactions containing 5 mM 4-pyrimidinone base and 75 mM ribose. Reactions were prepared on a watch glass and placed in an oven at 85˚C for 2 hr. Reaction mixtures were rehydrated in dH_2O and products were analyzed according to standard HPLC gradient conditions. Reactions were at pH 5.0 (solid, 0 mM HCl), pH 2.4 (dashed, 5 mM HCl), and pH 2.1 (dotted, 10 mM HCl).

Uracil was not observed to produce nucleosides when dried in the presence of ribose, which is consistent with the previous reports of Orgel and co-workers (Fuller, 1972). As discussed in preceding chapters, the reactivity of the pyrimidine bases with ribose is apparently not favorable because the delocalization of charge on N1 of the pyrimidine bases, which must initiate a nucleophilic attack toward the aldehydic carbon of ribose, poses a substantial kinetic barrier to pyrimidine glycosidic bond formation (Ingar, 2003). Based upon the successful production of nucleoside with 2-pyrimidinone, and presumably with 4-pyrimidinone, it was hypothesized that 5-azauracil and/or 6-azauracil (Figure 5.2) might also be able to form nucleosides in a similar heating and drying reaction with ribose. In Figure 5.4 HPLC chromatographs demonstrate that reaction products are formed when 5-azauracil is dried in the presence of ribose, both

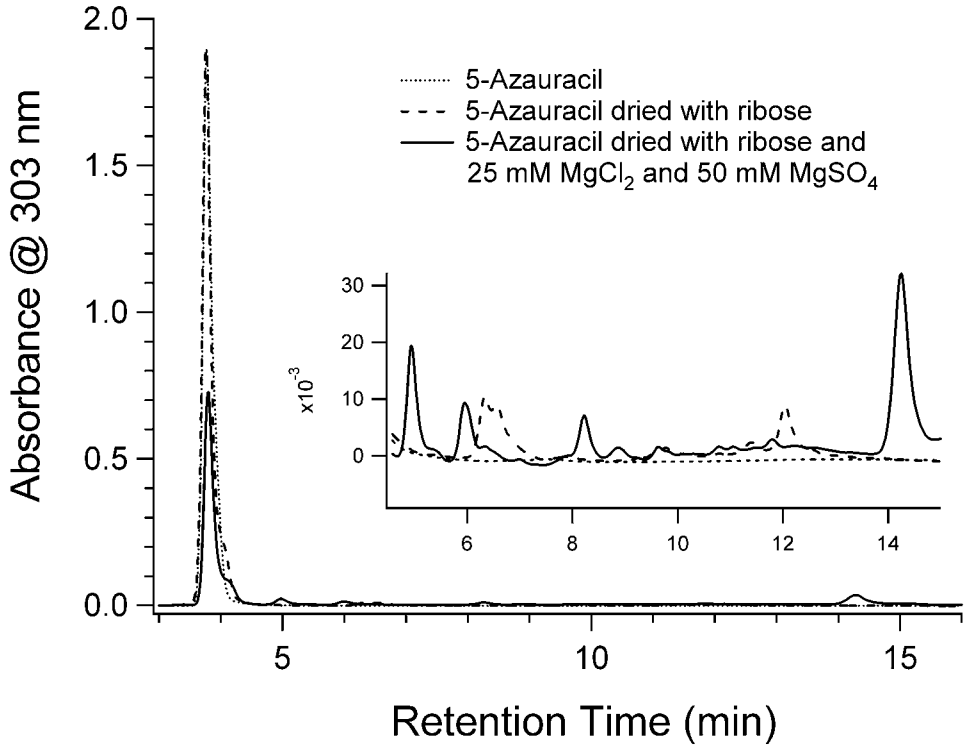

Figure 5.4: HPLC chromatographs of heating and drying reactions of 5-azauracil with and without ribose. Reaction mixtures contained 5 mM 5-azauracil base and 75 mM ribose, if present. Reactions were prepared on a watch glass and placed in an oven at 85˚C for 2 hr. Reaction mixtures were rehydrated in dH_2O and the products were analyzed according to standard HPLC gradient conditions. Dotted line is the 5-azauracil control sample, dashed line is the 5-azauracil reaction with ribose sugar, and the solid line is the 5-azauracil reaction with ribose in the presence of 25 mM $MgCl_2$ and 50 mM $MgSO_4$. Inset: Magnified region shows the appearance of new peaks that suggest the formation of nucleoside products.

with and without the addition of magnesium salts. These products have HPLC column retention times that are very similar to those of other pyrimidine nucleoside standards (an authentic standard for the possible 5-azauracil nucleosides was not available at the time of these experiments). The integrated intensity of these HPLC peaks strongly suggests that 5-azauracil forms nucleosides with ribose at about one-half the yield of 2-pyrimidinone. A similar reaction with 6-azauracil and ribose did not show any evidence of nucleoside formation (data not shown), which is consistent with previous attempts

made by Miller and co-workers to produce nucleosides with 6-azauracil under plausible prebiotic conditions (Kolb, 1994).

5.3 Can Other Sugars Form Nucleosides with 2-Pyrimidinone by the Same Reaction?

It is generally accepted that many sugars would have been present along with ribose on the prebiotic Earth. Furthermore, it is believed that ribose would have only been a minor component of this mixture (Decker, 1982). Thus, one of the great mysteries surrounding the origin of life is how and why nature chose ribose for the backbone of RNA. A number of studies have clearly demonstrated that other sugars can be used to make RNA-like molecules that form duplexes with base pairing stabilities that are comparable to RNA (Eschenmoser, 1999). One possible reason for the selection of ribose is that the bases were able to form nucleosides with ribose more easily than with other sugars. Another possibility is that the first RNA-like polymers used another sugar (or even a mixture of sugars), and that the incorporation of ribose was the result of molecular evolution.

To test if the nucleoside formation reaction described in this thesis for 2-pyrimidinone is particular to ribose, the same reaction was carried out using four other sugars. These four sugars were the three other D-pentose sugars and one hexose, D-glucose (Figure 5.5). Reaction mixtures were prepared with 5 mM 2-pyrimidinone base and 75 mM of each sugar. Reactions were prepared on watch glasses and placed in an oven at 85˚C for 2 hr. HPLC analysis of each sample was performed as described above. As shown in Figure 5.6, reactions with all three of the other pentoses and glucose resulted in products with HPLC column retention times consistent with nucleoside formation. Additionally, the integrated intensity of the HPLC peaks suggests that the yield of nucleosides with the 2-pyrimidinone base is roughly equal for all sugars tested. Thus, if the peaks observed in the chromatographs shown in Figure 5.6 are verified to be

nucleosides, then the 2-pyrimidinone nucleoside-formation reaction is not particular to ribose. In regards to the RNA world hypothesis, if the nucleoside-formation reaction is ultimately proven to work just as well with sugars other than ribose, then this would indicate that ribose was selected by evolution for its structural and functional attributes rather than for its ease in the synthesis of nucleosides.

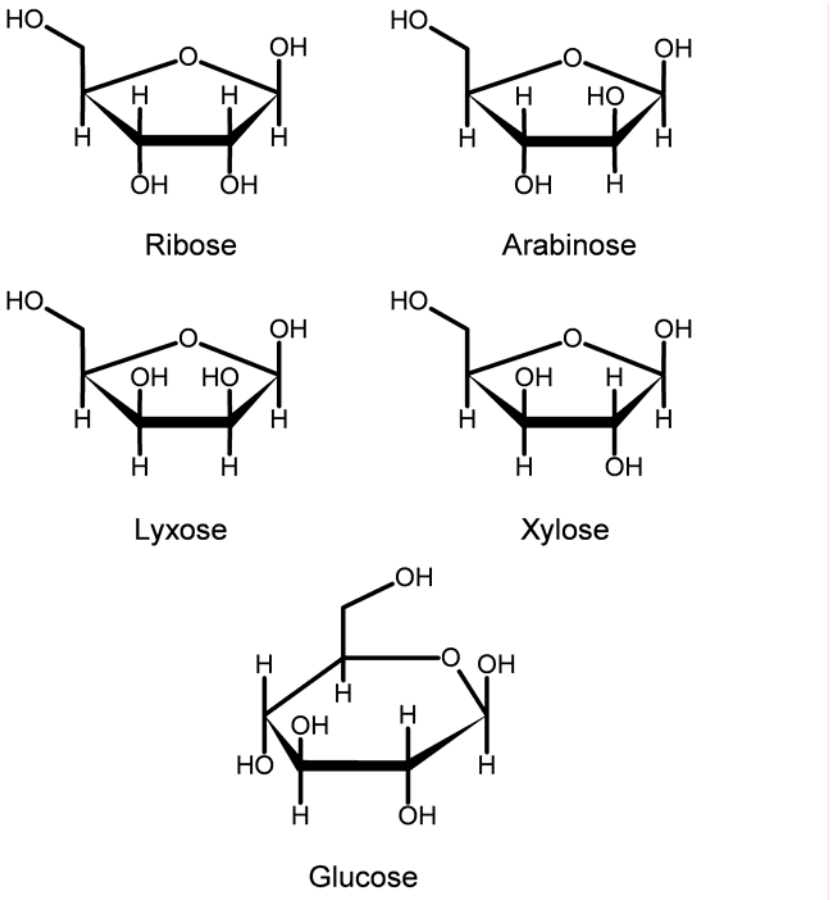

Figure 5.5: Structures of the five sugars tested in the 2-pyrimidinone nucleoside-formation reaction.

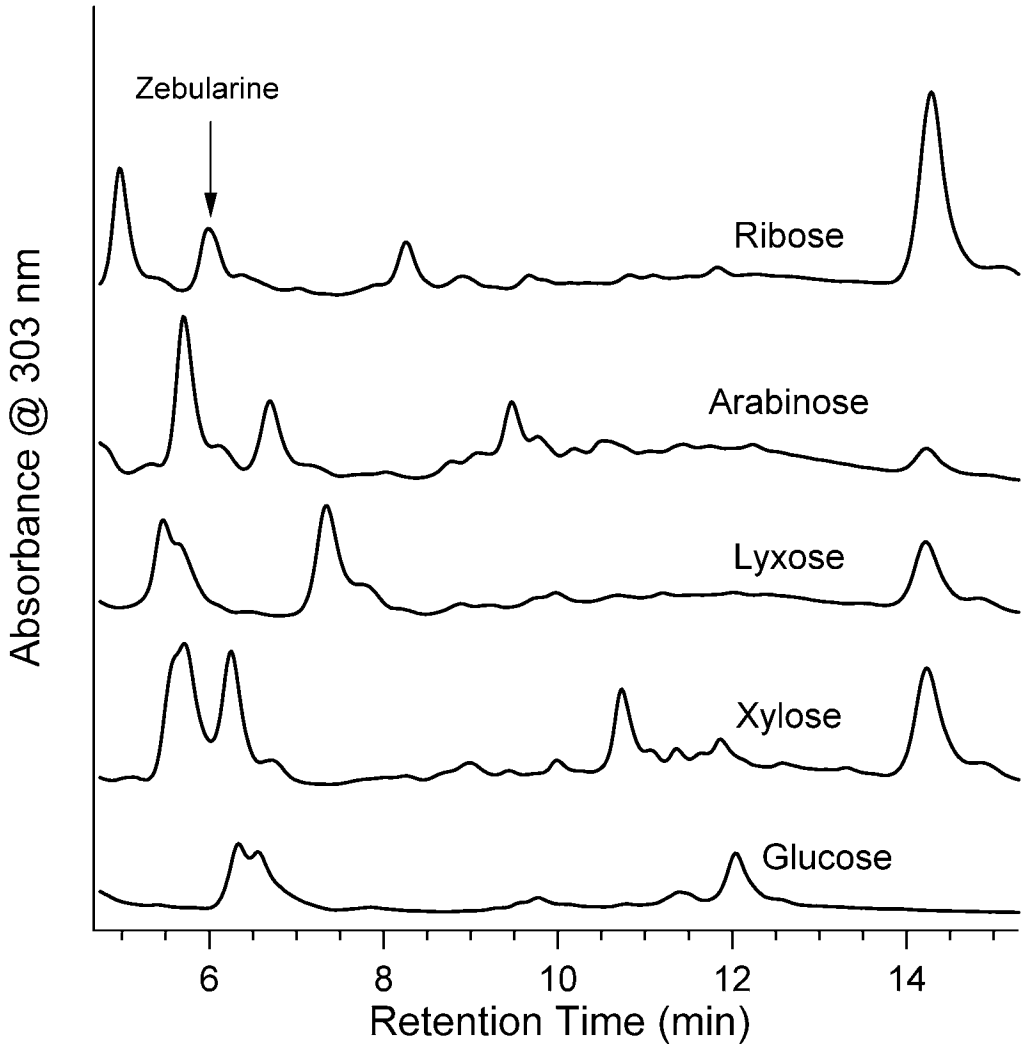

Figure 5.6: HPLC chromatographs of products resulting from reactions of 2-pyrimidinone with various sugars. Each reaction contained 5 mM 2-pyrimidinone base and 75 mM of the sugar indicated. Reactions were prepared on a watch glass and placed at 85˚C for 2 hr. Reaction mixtures were rehydrated in dH_2O, and the product mixtures were analyzed according to standard HPLC gradient conditions.

6. CONCLUSIONS

The primary objective of the work described in this thesis was to determine if the pyrimidine base 2-pyrimidinone would spontaneously form nucleosides with the sugar ribose when heated and dried from an aqueous solution. This work was motivated by the long-standing problem regarding how nucleosides were synthesized during the early stages of life, some 3.8 billion years ago. While the purine bases reportedly form nucleosides with ribose in a plausible prebiotic reaction, nucleoside formation with the pyrimidine bases has not been reported. The lack of a mechanism for the plausible prebiotic synthesis for the pyrimidine bases is considered by many researchers to be one of the most significant problems regarding most contemporary theories for the origin of life.

The discovery described in this thesis that 2-pyrimidinone will spontaneously form nucleosides with ribose in a plausible prebiotic reaction has significant implications regarding the origin of the first RNA-like polymers. It now seems feasible that both purine and pyrimidine nucleosides could have formed on the prebiotic Earth without the aid of protein enzymes. However, the fact remains that the pyrimidine bases presently found in RNA do not spontaneously form nucleosides in a simple heating and drying reaction. Together, these observations suggest that the pyrimidine bases in the ancestral RNA-like polymer were different from those of contemporary life, and that the pyrimidine bases of the original nucleosides were either modified to uridine and cytidine after the glycosidic bond had been formed, or the original pyrimidine bases were eventually replaced by uracil and cytosine. 2-pyrimidinone is a good candidate for a nucleoside base that could have been modified to uracil or cytosine after nucleoside formation, as this modification only requires oxidation or amination, respectively, at the C4 position of the pyrimidine ring. The preliminary results presented here regarding the

formation of nucleosides with 5-azauracil suggest an even more direct solution to the pyrimidine nucleoside problem, as it is very likely that 5-azauracil would have base pairing properties similar to that of uracil. The origin of cytidine, or cytidine-like, nucleosides still needs to be determined.

The results presented in this thesis have also demonstrated that 2-pyrimidinone spontaneously forms more than one nucleoside anomer with ribose, and that nucleosides can be formed with sugars other than ribose in similar yields. These results suggest that β-furanosyl nucleosides were not selected simply because they are formed in much greater yield, nor that ribose was selected from a prebiotic mixture of sugars because it is inherently better suited than other sugars for nucleoside formation. Rather, the results presented in this thesis suggest that ribose, and the β-furanose ribosides in particular, were selected because of their functional superiority to other nucleosides.

APPENDIX A

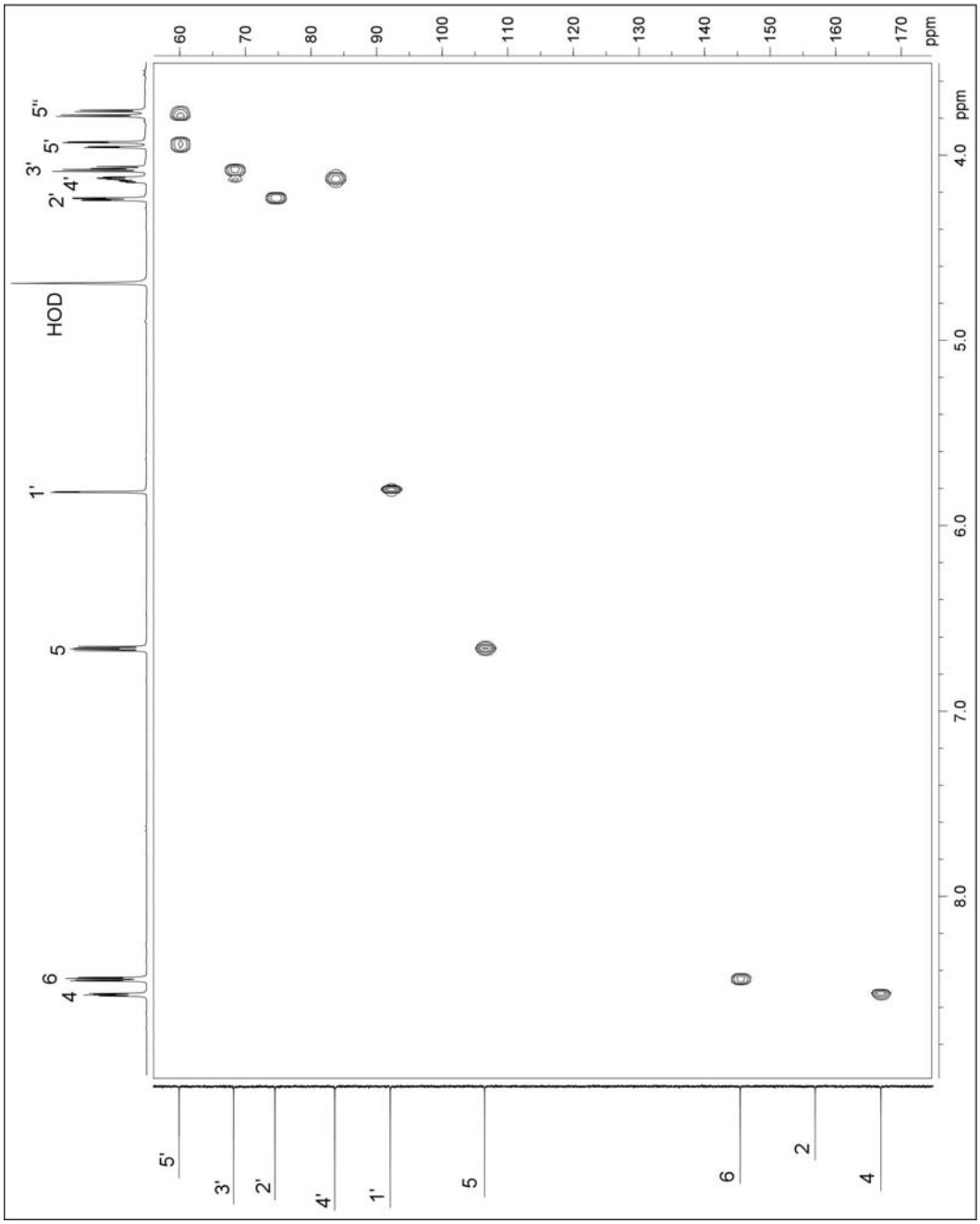

Figure A.1 HSQC spectrum of the zebularine standard. Spectrum recorded on a Bruker DRX500 spectrometer in D_2O at 25 ºC.

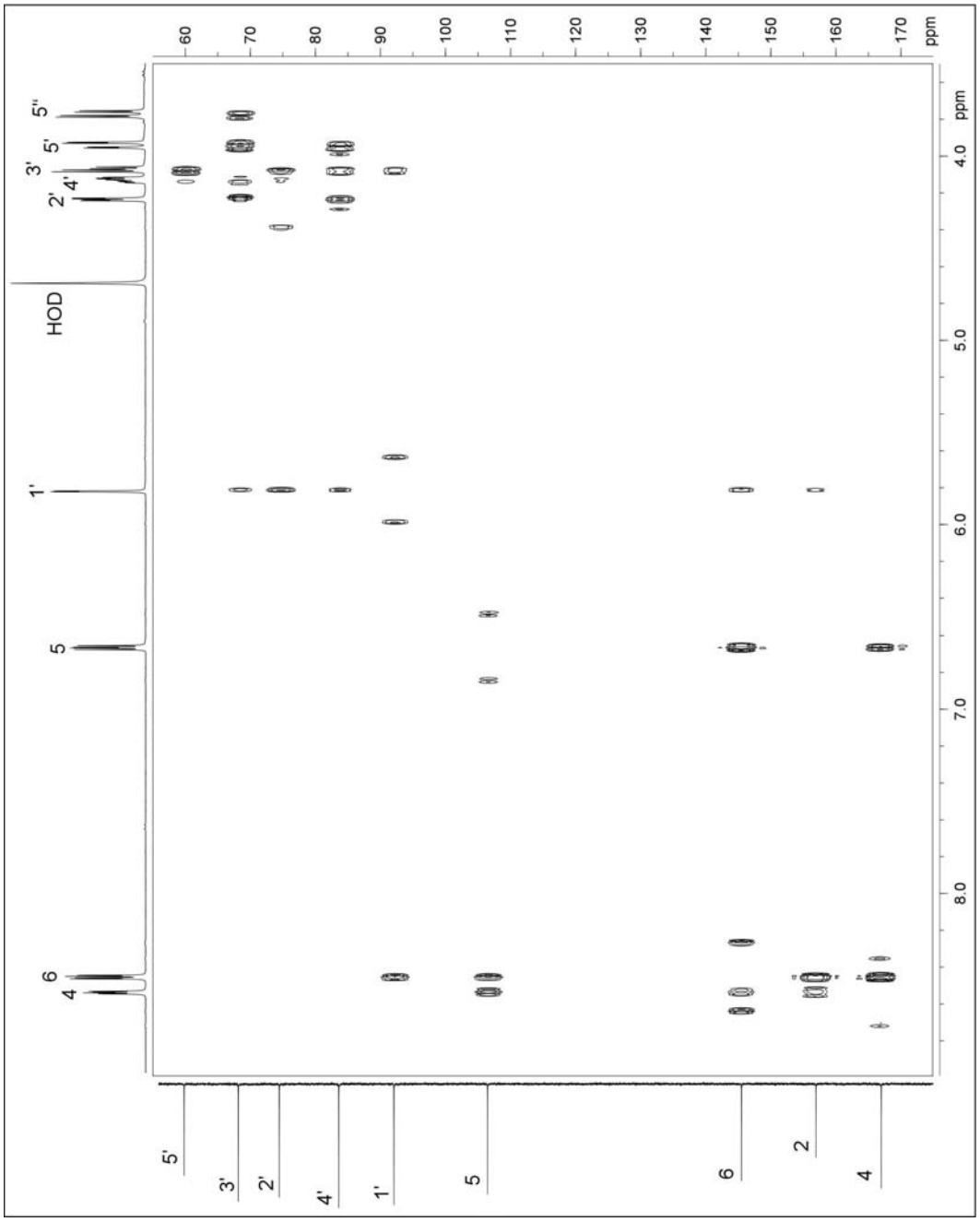

Figure A.2 HMBC spectrum of the zebularine standard. Spectrum recorded on a Bruker DRX500 spectrometer in D_2O at 25 ºC.

REFERENCES

Albert, A. and Phillips, J. N. (1956). "Ionization Constants of Heterocyclic Substances. Part II. Hydroxy-Derivatives of Nitrogenous Six-Membered Ring-Compounds." *J Chem Soc*: 1294-1304.

Bolli, M., Micura, R. and Eschenmoser, A. (1997). "Pyranosyl-RNA: Chiroselective Self-Assembly of Base Sequences by Ligative Oligomerization of Tetra Nucleotide-2′,3′-Cyclophosphates (with a Commentary Concerning the Origin of Biomolecular Homochirality)." *Chem Biol* **4**: 309-320.

Bengston, S. (1994). Early Life on Earth. New York, Columbia University.

Carey, F. A. (1996). Organic Chemistry. New York, The McGraw-Hill Companies, Inc.

Crick, F. H. C. (1968). "The Origin of the Genetic Code." *J Mol Biol* **38**: 367-379.

Deamer, D. W. and Fleischaker, G. R., Eds. (1994). Origins of Life: The Central Concepts. Boston, Jones and Bartlett.

Decker, P., Schweer, P. and Pohlmann, R. (1982). "Identification of Formose Sugars, Presumable Prebiotic Metabolites, Using Capillary Gas Chromatography/Gas Chromatography-Mass Spectroscopy of N-Butoxime Trifluoroacetates on Ov-225." *J Chromatogr* **244**: 281-291.

Driscoll, J. S., Marquez, V. E., Plowman, J., Liu, P. S., Kelley, J. A. and Barchi, J. J. (1991). "Antitumor Properties of 2(1H)-Pyrimidinone Riboside (Zebularine) and Its Fluorinated Analogs." *J Med Chem* **34**: 3280-3284.

Ellington, A. and Szostak, J. (1990). "In Vitro Selection of RNA Molecules That Bind Specific Ligands." *Nature* **346**: 818-822.

Eschenmoser, A. (1999). "Chemical Etiology of Nucleic Acid Structure." *Science* **284**: 2118-2124.

Fox, S. W. and Harada, K. (1958). "Thermal Copolymerization of Amino Acids to a Product Resembling Protein." *Science* **128**: 1214-1214.

Fox, S. W., Harada, K. and Kendrick, J. (1959). "Production of Spherules from Synthetic Proteinoid and Hot Water." *Science* **129**: 1221-1223.

Fuller, W. D., Sanchez, R. A. and Orgel, L. E. (1972). "Studies in Prebiotic Synthesis: VII. Solid-State Synthesis of Purine Nucleosides." *J Mol Evol* **1**: 249-257.

Gilbert, W. (1986). "The RNA World." *Nature* **319**: 618-618.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983). "The RNA Moiety of Ribonuclease P Is the Catalytic Subunit of the Enzyme." *Cell* **35**: 849-857.

Ingar, A.-A., Luke, R. W. A., Hayter, B. R. and Sutherland, J. D. (2003). "Synthesis of Cytidine Ribonucleotides by Stepwise Assembly of the Heterocycle on a Sugar Phosphate." *ChemBioChem* **4**: 504-507.

Iocono, J. A., Gildea, B. and McLaughlin, L. W. (1990). "Mild Acid Hydrolysis of 2-Pyrimidinone-Containing DNA Fragments Generates Apurinic/Apyrimidinic Sites." *Tetrahedron Lett* **31**: 175-178.

Joyce, G. F. (1999a). Appendix 3: Reactions Catalyzed by RNA and DNA Enzymes. The RNA World, Second Edition: The Nature of Modern RNA Suggests a Prebiotic RNA World. Gesteland, R. F. and Atkins, J. F. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press**:** 687-690.

Joyce, G. F. and Orgel, L. E. (1999b). Prospects for Understanding the Origin of the RNA World. The RNA World, Second Edition: The Nature of Modern RNA Suggests a Prebiotic RNA World. Gesteland, R. F. and Atkins, J. F., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press**:** 49-77.

Katritzky, A. R. (2000). Handbook of Hetercyclic Chemistry. New York, Pergamon.

Keefe, A. D. and Miller, S. L. (1995). "Are Polyphosphates or Phosphate Esters Prebiotic Reagents?" *J Mol Evol* **41**: 693-702.

Kett, W. C., Batley, M. and Redmond, J. W. (1997). "Heterocyclic Derivatives of Sugars: An NMR Study of the Formation of 1-Glycosyl-3,5-Dimethyl-1 H-Pyrazoles from Hydrazones." *Carbohydr Res* **299**: 129-141.

Kolb, V. M., Dworkin, J. P. and Miller, S. L. (1994). "Alternative Bases in the RNA World: The Prebiotic Synthesis of Urazole and Its Ribosides." *J Mol Evol* **38**: 549-557.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982). "Self-Splicing RNA: Autoexcision and Autocyclization of the Ribosomal RNA Intervening Sequence of *Tetrahymena*." *Cell* **31**: 147-157.

Larralde, R., Robertson, M. P. and Miller, S. L. (1995). "Rates of Decomposition of Ribose and Other Sugars: Implications for Chemical Evolution." *PNAS* **92**: 8158-8160.

Miller, S. L. (1953). "A Production of Amino Acids under Possible Primitive Earth Conditions." *Science* **117**: 528-529.

Noller, H. F., Hoffarth, V. and Zimniak, L. (1992). "Unusual Resistance of Peptidyl Transferase to Protein Extraction Procedures." *Science* **256**: 1416-1419.

Orgel, L. E. (1968). "Evolution of the Genetic Apparatus." *J Mol Biol* **38**: 381-393.

Saewan, N., Crowe, M. A., Helliwell, M., Raftery, J., Chantrapromma, K. and Sutherland, J. D. (2005). "Exploratory Studies to Investigate a Linked Prebiotic Origin of RNA and Coded Peptides - 4th Communication - Further Observations Concerning Pyrimidine Nucleoside Synthesis by Stepwise Nucleobase Assembly." *Chem Biodivers* **2**: 66-83.

Saladino, R., Ciambecchini, U., Crestini, C., Costanzo, G., Negri, R. and Di Mauro, E. (2003). "One-Pot TiO_2-Catalyzed Synthesis of Nucleic Bases and Acyclonucleosides from Formamide: Implications for the Origin of Life." *ChemBioChem* **4**: 514-521.

Sanchez, R. A. and Orgel, L. E. (1970). "Studies in Prebiotic Synthesis. V. Synthesis and Photoanomerization of Pyrimidine Nucleosides." *J Mol Biol* **47**: 531-543.

Schopf, J. W. (1999). Cradle of Life: The Discovery of Earth's Earliest Fossils. Princeton, Princeton University Press.

Shapiro, R. and Klein, R. S. (1966). "Deamination of Cytidine and Cytosine by Acidic Buffer Solutions. Mutagenic Implications." *Biochemistry* **5**: 2358-2362.

Tran-Dinh, S., Neumann, J. M., Thiery, J. M., Tam, H. D., Igolen, J. and Guschlbauer, W. (1977). "Configuration and Conformation of the α- and β-Anomers of C-Nucleosides by Proton Magnetic Resonance Spectroscopy: New Criterion for Determination of α- and β-Anomers." *J Am Chem Soc* **99**: 3267-3273.

Tuerk, C. and Gold, L. (1990). "Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase." *Science* **249**: 505-510.

Wani, M. C., Nicholas, A. W. and Wall, M. E. (1986). "Plant Antitumor Agents. 23. Synthesis and Antileukemic Activity of Camptothecin Analogs." *J Med Chem* **29**: 2358-2363.

Woese, C. R. (1967). The Genetic Code: The Molecular Basis for Genetic Expression. New York, Harper and Row.

Zubay, G. and Mui, T. (2001). "Prebiotic Synthesis of Nucleosides." *Origins Life Evol B* **31**: 87-102.