

## Prebiotic Synthesis of Diaminopyrimidine and Thiocytosine

Michael P. Robertson,\* Matthew Levy, Stanley L. Miller

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0317, USA

Received: 12 December 1995 / Accepted: 29 June 1996

**Abstract.** The reaction of guanidine hydrochloride with cyanoacetaldehyde gives high yields (40–85%) of 2,4-diaminopyrimidine under the concentrated conditions of a drying lagoon model of prebiotic synthesis, in contrast to the low yields previously obtained under more dilute conditions. The prebiotic source of cyanoacetaldehyde, cyanoacetylene, is produced from electric discharges under reducing conditions. The effect of pH and concentration of guanidine hydrochloride on the rate of synthesis and yield of diaminopyrimidine were investigated, as well as the hydrolysis of diaminopyrimidine to cytosine, isocytosine, and uracil. Thiourea also reacts with cyanoacetaldehyde to give 2-thiocytosine, but the pyrimidine yields are much lower than with guanidine hydrochloride or urea. Thiocytosine hydrolyzes to thio-uracil and cytosine and then to uracil. This synthesis would have been a significant prebiotic source of 2-thiopyrimidines and 5-substituted derivatives of thio-uracil, many of which occur in tRNA. The applicability of these results to the drying lagoon model of prebiotic synthesis was tested by dry-down experiments where dilute solutions of cyanoacetaldehyde, guanidine hydrochloride, and 0.5 M NaCl were evaporated over varying periods of time. The yields of diaminopyrimidine varied from 1 to 7%. These results show that drying lagoons and beaches may have been major sites of prebiotic syntheses.

**Key words:** Pyrimidine synthesis — Cyanoacetaldehyde — Guanidine — Thiourea—Drying lagoons

### Introduction

Until recently, the prebiotic synthesis of pyrimidines has been considered poor. The first experiments used cyanoacetylene and cyanate to produce cytosine, but the concentrations of cyanate needed were rather high (0.1 M) (Ferris et al. 1968). An overlooked experiment in this paper used 1 M cyanoacetylene and 1 M urea to produce cytosine in 4.8% yield.

Cyanoacetylene is produced in substantial yield from electric discharges acting on CH<sub>4</sub> + N<sub>2</sub> mixtures (Sanchez et al. 1966), but it hydrolyzes rather rapidly to cyanoacetaldehyde (Ferris et al. 1970) especially in the presence of H<sub>2</sub>S. Cyanoacetaldehyde was shown to be a likely product of electric discharges as the Strecker precursor to aspartic acid (Miller 1957).

Ferris et al. (1974) showed a 0.1–2% yield of diaminopyrimidine from cyanoacetaldehyde and 10<sup>-3</sup> to 0.1 M guanidine. However, there was no detectable cytosine (or uracil) produced from cyanoacetaldehyde and dilute urea solution.

We have recently shown that cytosine can be efficiently synthesized from concentrated solutions of urea and cyanoacetaldehyde with yields as high as 53% (Robertson and Miller 1995a). Since cytosine, isocytosine, and uracil are produced by hydrolysis of diaminopyrimidine, we decided to investigate the reaction of cyanoacetaldehyde with guanidine at high guanidine concentrations because most salts of guanidine are very soluble.

**Abbreviations:** DAP, diaminopyrimidine; Cy, cyanoacetaldehyde; C, cytosine; IC, isocytosine; U, uracil; G, guanidine; GH<sup>+</sup>, guanidinium; G·HCl, guanidinium chloride; SC, 2-thiocytosine; SU, 2-thiouracil

\* *Present address:* Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

*Correspondence to:* S.L. Miller

Diaminopyrimidine is of additional interest since it has been proposed as an alternative base pair with xanthine (Piccirilli et al. 1990), and it has been incorporated into DNA and RNA with DNA and RNA polymerases. We find that yields as high as 85% of diaminopyrimidine can be obtained with the drying lagoon model of prebiotic synthesis. We have also done a preliminary investigation of the rate of hydrolysis of diaminopyrimidine and shown that it hydrolyzes to both cytosine and isocytosine, and then to uracil.

The analogous reaction with thiourea was briefly reported to give low yields of 2-thiocytosine (Robertson and Miller 1995a). Although the yields are low, this synthesis may be of considerable significance since there are a large number of 5-substituted thiouracils in tRNA, generally at the wobble position. We have further investigated the thiourea reaction and the hydrolytic stability of thiocytosine and thiouracil.

## Materials and Methods

The diaminopyrimidine, isocytosine, cytosine, and uracil were obtained from Sigma. The guanidine hydrochloride was from Sigma or Eastman and used after recrystallization from H<sub>2</sub>O. The cyanoacetaldehyde was prepared from isoxazole (Ferris et al. 1968). The preparation of the samples and the HPLC system used have been previously described (Robertson and Miller 1995a). The DAP was identified by elution times on the reversed phase and cation HPLC columns. The identification was confirmed by a spectrophotometric titration of the DAP peak collected from the reversed-phase column.

## Results

### Diaminopyrimidine

The first variable investigated in diaminopyrimidine synthesis was the effect of pH on the yields using 1 *m* (molal<sup>1</sup>) G · HCl (Fig. 1). It is apparent that the yield increases with increasing pH although there is a plateau between pH 7 and 9. The yields at infinite time range from 4% at pH 3 to 53% at pH 12.

The rate as a function of guanidine concentration was investigated at pH 7.6 (Fig. 2). The rate is approximately proportional to the molality of guanidine and the pseudo-first-order rate constants are given by the equation

$$k \text{ (m}^{-1}\text{s}^{-1}\text{)} = 9.29 \times 10^{-7} \text{ (100}^\circ\text{C, pH 7.6)}$$

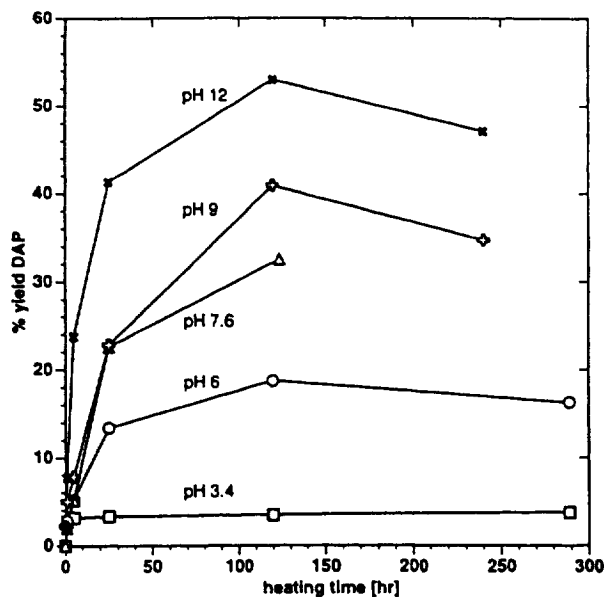


Fig. 1. The yield of DAP with heating time at 100°C and 1 *m* GH<sup>+</sup>. The samples were buffered with 0.02 *M* succinate, phosphate, or borate.

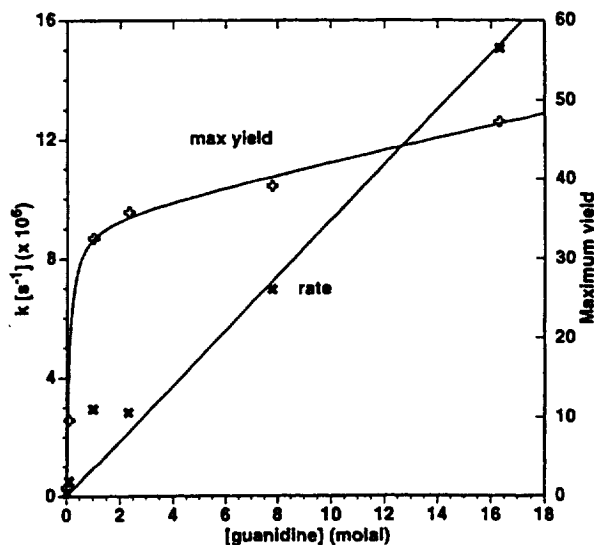
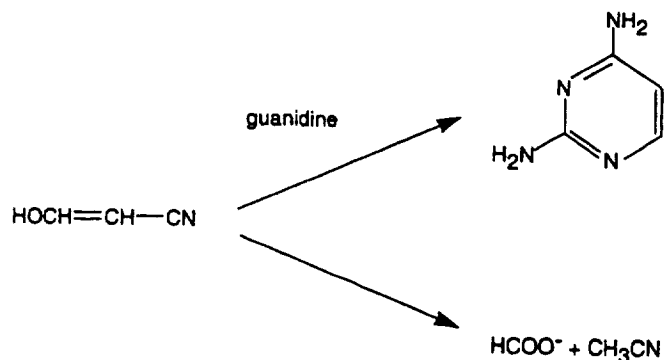


Fig. 2. Rate of formation and maximum yield of diaminopyrimidine as a function of GH<sup>+</sup> molality at pH 7.6, 0.04 *M* phosphate, 100°C. The pH was essentially constant in the rate measurement samples (5-h heating times) but the pH rose to 9 with the longer heating times of the maximum yield samples.

This rate is close to that of cytosine synthesis from urea ( $9.48 \times 10^{-7} \text{ s}^{-1}$  with 1 *m* urea) (Robertson and Miller 1995a).

The maximum yield is also shown on Fig. 2. Not shown is the yield of 85% at 100° using saturated G · HCl (62 *m*), which makes this a very efficient prebiotic synthesis. The reason for the lowered yield at low guanidine concentrations are the competing reactions of diaminopyrimidine formation and cyanoacetaldehyde decomposition as previously described (Ferris et al. 1974). The temperature (*T*) coefficient of this reaction

<sup>1</sup> Abbreviated *m*. We use molal (moles per 1000 g H<sub>2</sub>O) instead of molar (*M*) (moles per liter) because molal concentrations are independent of temperature and molar solutions are not. Molal G · HCl solutions can be converted to molar by the equation  $m/M = 0.997 - 0.0593m + 0.00166m^2$ .



was measured at pH 7 with 1 *m* guanidine. The Arrhenius curve is shown in Fig. 3. The rate is given by

$$\log k \text{ (} m^{-1} \text{ s}^{-1}\text{)} = 9.02 - 5,478/T$$

The heat of activation is 25.1 kcal. The half-life at 25°C is 50 years for 1 *m* G · HCl and 2.3 years in saturated G · HCl (23 *m*).

#### Solubility

Most salts of guanidine are very soluble. Examples at 30°C are guanidine sulfate (12.5 *m*), guanidine nitrate (1.6 *m*), guanidine thiocyanate (13.5 *m*), and guanidine carbonate (2.8 *m*) (Elvers et al. 1989). Guanidine hydrochloride is among the most soluble. Elvers et al. (1989) report 23.9 *m* at 30°C and 47.1 *m* at 90°C; and Kirk and Othmer (1951) report 22.5 *m* at 20°C and 33.5 *m* at 55°C, but no source for these data is cited. We therefore measured the solubility at several temperatures by approaching equilibrium from both the undersaturated and oversaturated sides, using sealed glass tubes and a constant temperature bath. An aliquot of the saturated solution was weighed, dried in a desiccator, and weighed again.

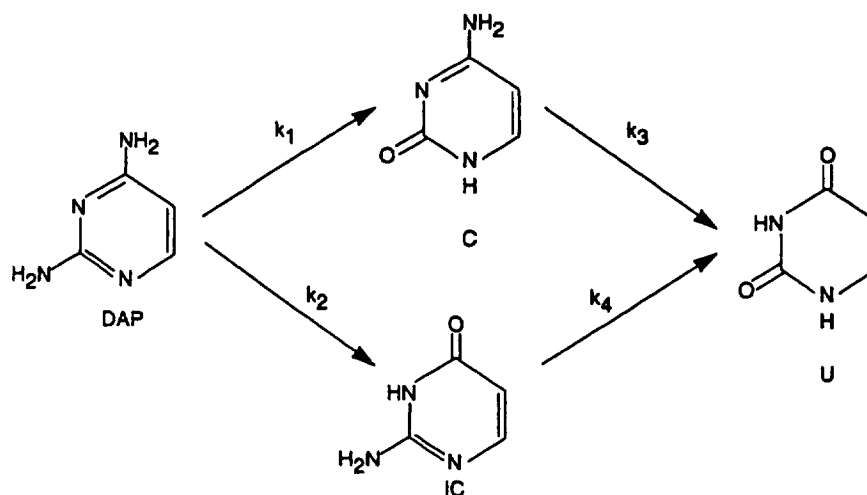
The 99°C solubility was determined by preparing a 60.0 *m* solution and determining the temperature (99.0°) where the G · HCl disappeared on heating and reappeared on cooling. The solubilities are 17.2 *m* at 0°C and 22.9 *m* at 25°C. These solubilities can be expressed in terms of the mole fraction,  $X_2$ , by the equation:

$$\log X_2 = -1.293 - \frac{42.81}{T} + 3.031 \times 10^{-3} T$$

This includes the melting point of 184°C where  $X_2 = 1$  and the reported solubilities at 20°, 30°, and 55°C.

#### Hydrolysis of Diaminopyrimidine

The hydrolysis of diaminopyrimidine was previously reported, on the basis of paper chromatographic evidence, to give cytosine, and to give uracil from the hydrolysis of cytosine (Ferris et al. 1974). Since the HPLC system permits the detection of small amounts of isocytosine, we decided to examine the hydrolysis of diaminopyrimidine in more detail. The scheme is:



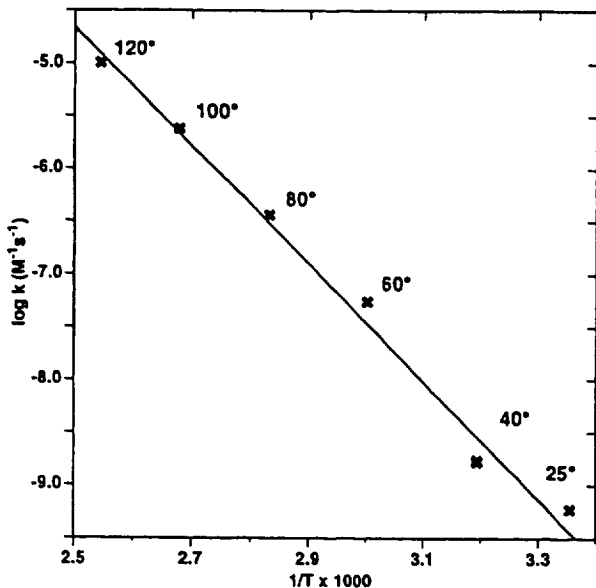


Fig. 3. Arrhenius plot for the formation of diaminopyrimidine. The second-order rate constants were measured using 1.0 *m* GHCl.

The kinetic equations are:

$$-\frac{d[\text{DAP}]}{dt} = (k_1 + k_2)[\text{DAP}]$$

$$+\frac{d[\text{C}]}{dt} = k_1[\text{DAP}] - k_3[\text{C}]$$

$$+\frac{d[\text{IC}]}{dt} = k_2[\text{DAP}] - k_4[\text{IC}]$$

$$+\frac{d[\text{U}]}{dt} = k_3[\text{C}] + k_4[\text{IC}]$$

Taking  $[\text{DAP}]_0$  as the initial  $[\text{DAP}]$  and  $\text{C} = \text{IC} = \text{U} = 0$ , the solution to these equations is (Frost and Pearson 1961):

$$[\text{DAP}] = [\text{DAP}]_0 e^{-(k_1 + k_2)t}$$

$$[\text{C}] = \frac{[\text{DAP}]_0 k_1}{k_3 - (k_1 + k_2)} [e^{-(k_1 + k_2)t} - e^{-k_3 t}]$$

$$[\text{IC}] = \frac{[\text{DAP}]_0 k_2}{k_4 - (k_1 + k_2)} [e^{-(k_1 + k_2)t} - e^{-k_4 t}]$$

$$[\text{U}] = [\text{DAP}]_0 - [\text{DAP}] - [\text{C}] - [\text{IC}]$$

The values of  $k_3$  and  $k_4$  at pH 7 and 100°C were determined by measuring the rate of hydrolysis of C and IC to U, giving:

$$k_3 = 5.36 \times 10^{-7} \text{ s}^{-1}$$

$$k_4 = 3.78 \times 10^{-7} \text{ s}^{-1}$$

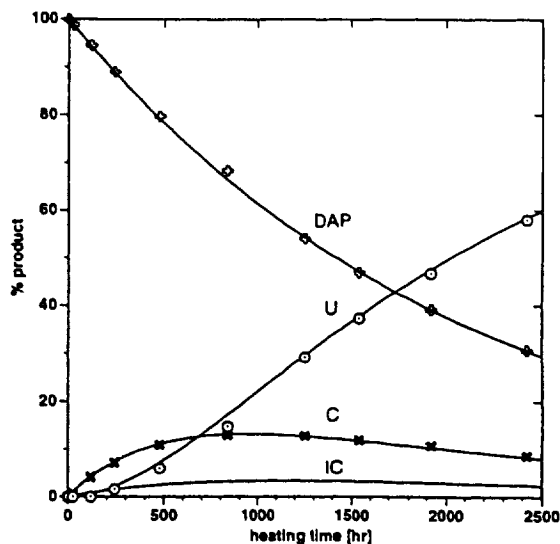


Fig. 4. Time course of diaminopyrimidine hydrolysis at pH 7.0 (0.02 *M* phosphate), 100°C. The lines are calculated from the rate constants and equations in the text.

The value of  $k_1 + k_2$  was determined by the rate of loss of DAP. The ratio of  $k_1/k_2$  is given by the initial rate of production of C and IC on hydrolysis of DAP. This gives:

$$k_1 = 1.13 \times 10^{-7} \text{ s}^{-1}$$

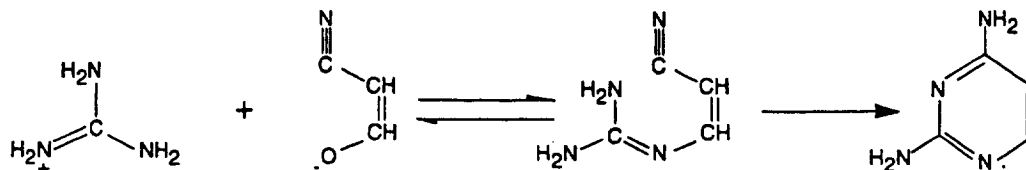
$$k_2 = 2.26 \times 10^{-8} \text{ s}^{-1}$$

With these rate constants available, the experimental time course of the hydrolysis of DAP to U can be compared with the differential equations as shown in Fig. 4. The agreement of the experimental points with the equations is good. Because of the difficulty of measuring IC by HPLC only the initial IC concentration was determined. The calculated ratio of IC/C is approximately 0.28 during most of the hydrolysis.

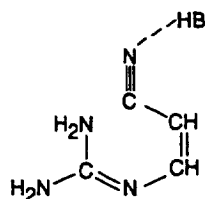
#### pH-Rate Profile

The pH-rate profile was investigated using 1 *M* guanidinium hydrochloride with the pH maintained by a series of buffers. The reactions were stopped after only a few percent reaction to minimize the pH drift due to the decomposition of guanidine.

The second-order rate constants are shown in Fig. 5. It is a rather complex pH-rate profile showing an acid-catalyzed leg, a plateau between pH 5 and 7, a second plateau between pH 10 and 11, followed by a falloff in the rate. We interpret the pH 8–12 part of the curve as the reactions:



The falloff in the rate at pH 11 is due to the ionization of guanidinium. This  $pK$  is estimated to be 11.0 at 100°C from the  $pK = 13.65$  at 25°C and the heat of ionization of 18.2 kcal (Fabbrizzi et al. 1977). The falloff in rate below pH 9 may be due to protonation of the cyanoacetaldehyde. The pH region between 4 and 7 may be due to general acid and base catalysis of the ring closure of



with the acid-catalyzed leg being specific  $H^+$  catalysis of the ring closure. This interpretation is tentative because the  $pK$ s are not known and alternative pathways may be involved.

#### Thiocytosine and Thiouracil

The possibility of synthesizing 2-thiocytosine and 2-thiouracil by an analogous reaction with thiourea is attractive because these bases occur in tRNA, as do a number of 5-substituted 2-thiouracils (Limbach et al. 1994). These modified bases generally occur in the wobble position of the anticodon.

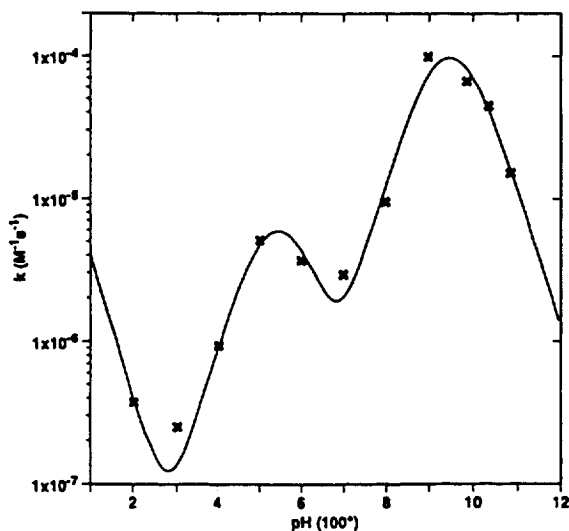
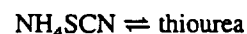


Fig. 5. The pH rate profile at 100°C for the synthesis of DAP with 1 *m* GHCl. The reaction proceeded to yields of 1–5% to avoid drift in pH. This drift was <0.1 pH unit except for pH 11 and 12 (pH 10.4 and 10.9 at 100°C), for which drift was 0.5.

There are two prebiotic sources of thiourea. The first is the addition of  $H_2S$  to cyanamide, both of which are likely prebiotic compounds (Lohrman 1972). The second source would be the solid reaction



which is analogous to the Wöhler urea synthesis from ammonium cyanate. The equilibrium is less favorable, with a maximum thiourea yield of ~25% (Reid 1963), compared to the nearly quantitative yield of urea with  $NH_4NCO$ . Quantitative yields of thiourea, however, are not needed. An advantage of thiourea over urea is that while it is more difficult to synthesize, it decomposes 244 times more slowly than urea (Shaw and Walker 1958). Thus, on balance, thiourea seems to be nearly as plausible a prebiotic reagent as urea. Although the solubility of thiourea (2.2 *m* at 25°C) is considerably less than urea at low temperatures, they become closer at 100°C where thiourea solubility is 52 *m* and urea is 120 *m*. The solubility in terms of the mole fraction ( $X_2$ ) is given by Shnidman (1933) as:

$$\log X_2 = 4.029 - \frac{1,621.6}{T}$$

We recently reported briefly the prebiotic synthesis of thiocytosine from cyanoacetaldehyde and 1 mol thiourea in 0.05% yield (Robertson and Miller 1995a), and we also showed that several 5-substituted uracils could be obtained by hydroxymethylation of SU followed by substitution of various nucleophiles (e.g., glycine) (Robertson and Miller 1995b).

Samples of  $10^{-3}$  *M* cyanoacetaldehyde and various concentrations of thiourea at pH 7 were heated in sealed tubes. Thiourea decomposes much more slowly to ammonium thiocyanate than urea decomposes to ammonium cyanate, but even so, the pH drifted slightly lower. It is to be noted that the HPLC chromatograms show peaks corresponding to thiocytosine and thiouracil, but the nuclear magnetic resonance (NMR) of the isolated thiocytosine peak does not correspond to thiocytosine. If the molar absorptivity of the unknown compound in the peak were  $10^4$   $M^{-1} \text{ cm}^{-1}$ , then the yield would be ~50%.

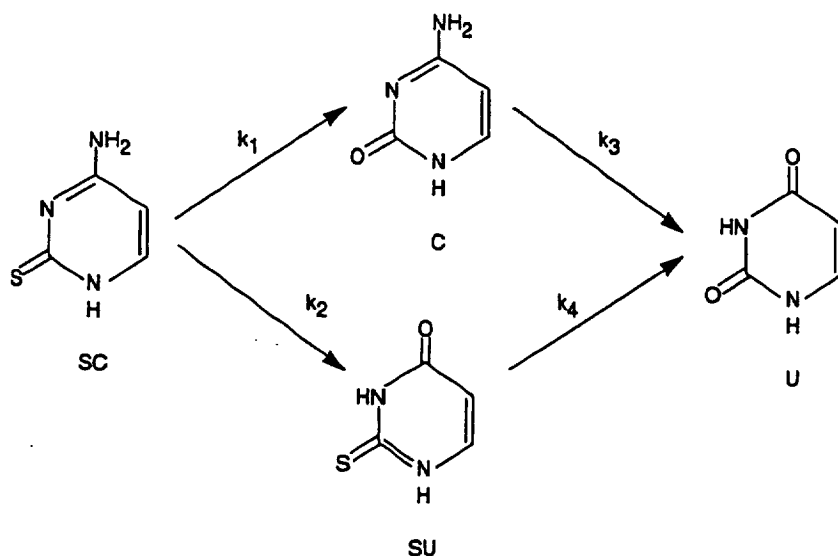
The thiocytosine was isolated by chromatography of the heated sample of 1 *M* thiourea and  $10^{-3}$  *M* cyanoacetaldehyde on a Dowex 50 ( $H^+$ ) column using 1 *M* HCl as eluent. The peak eluting from the Dowex 50 at the SC position gave an HPLC peak at the correct elution time and the NMR spectrum was identical to the known SC.

Using higher concentrations of thiourea presumably gave higher yields of SC but isolation was complicated by increasing amounts of side products, which greatly interfered with HPLC analysis. Several experiments using buffers of different pH or different temperatures did not greatly increase the yield of SC.

The reaction of cyanoacetaldehyde with 1, 25, and 100 *m*  $\text{NH}_4\text{SCN}$  did not give detectable yields of SC (<0.01%).

We also attempted to synthesize SC and SU by heating DAP at 100°C with 0.01 M  $\text{H}_2\text{S}$  at pH 7 for 20 days. There was no detectable yield of SC (<0.01%) or SU (<0.01%). Apparently the addition of  $\text{H}_2\text{O}$  to DAP is favored over the addition of  $\text{H}_2\text{S}$ .

The product of the thiourea reaction is 2-thiocytosine. Since most of the modified nucleosides in tRNA are 2-thiouracils, we investigated the rate of hydrolysis of 2-thiocytosine to 2-thiouracil. The kinetic scheme is shown:



At pH 7 and 100°C, the value of  $k_3$  is obtained from the DAP data,  $k_4$  from a separate experiment on the rate of hydrolysis of thiouracil to uracil, and  $k_1 + k_2 = 7.70 \times 10^{-7} \text{ s}^{-1}$  from the rate of loss of thiocytosine to SC, SU, and U. The ratio of  $k_1/k_2 = 0.097$  was obtained from the steady-state ratio of C and SU. The values of these rate constants are:

$$k_1 = 6.8 \times 10^{-8} \text{ s}^{-1}$$

$$k_2 = 7.0 \times 10^{-7} \text{ s}^{-1}$$

$$k_3 = 5.4 \times 10^{-7} \text{ s}^{-1}$$

$$k_4 = 2.9 \times 10^{-7} \text{ s}^{-1}$$

It should be emphasized that these numbers are approximate, but they show that SC hydrolyzes preferentially to SU. The steady-state ratio of SU/C is about 20, showing that SC is a good source for SU but a poor source for C. However, the SU eventually hydrolyzes to U as does C.

### Dry-Down Experiments

In order to better simulate a drying lagoon prebiotic synthesis, dilute solutions of cyanoacetaldehyde ( $\sim 1 \times 10^{-3} \text{ M}$ ) and guanidine  $\cdot$  HCl ( $10^{-2} \text{ M}$ ) were taken to dryness with open test tubes in heating blocks both in the presence and absence of 0.5 M NaCl. The yield of DAP for samples dried down at 100°C for 27 h was 4.0–7.2% with NaCl and nearly the same (2.6–7.4%) without. Longer heating of these 27-h samples did not substantially increase the yield. The variation in yields is due to the variable rates of evaporation in the heating blocks and the nonreproducible crystallization process. Samples dried down at 80°C for 13 days gave yields of 4.0–5.7% with NaCl and somewhat lower yields (0.6–1.9%) with water.

### Discussion

The yields of DAP in these experiments are in agreement with the yields reported by Ferris et al. (1974) at low

concentrations of guanidinium chloride. At higher G  $\cdot$  HCl concentrations the yields and rate of formation of DAP are approximately proportional to the molality of G  $\cdot$  HCl as with urea. The yield, however, reaches a maximum at around 50% because of the competing reaction of cyanoacetaldehyde decomposition. Since the yield is quite low at G  $\cdot$  HCl concentrations less than 1 mol, this synthesis is applicable only to drying lagoons and beaches and other places where the guanidine can be concentrated from seawater.

The best prebiotic synthesis of G occurs by the addition of  $\text{NH}_3$  to cyanamide, the cyanamide coming from the photolysis of HCN solutions (Lohrmann 1972). G is also produced in the polymerization of HCN (Lowe et al. 1963). The yields are not particularly high in these experiments, but the steady-state concentration of G  $\cdot$  HCl would be higher than indicated by the yields especially in drying lagoons and on beaches because of its stability.

Guanidinium nitrate, guanidinium thiocyanate, and other guanidinium salts are less soluble than  $G \cdot HCl$ , but since  $Cl^-$  would be the dominant anion, the solubility of guanidinium would be determined by the solubility of  $G \cdot HCl$ . This very soluble salt is also very stable to hydrolysis (Ferris et al. 1974), considerably more stable than urea. Thus the dilute guanidinium in the seawater would concentrate in the drying lagoons and beaches and precipitate out at about the same time as the urea but considerably after the  $NaCl$  and  $CaCl_2$ . Since cyanoacetaldehyde is nonvolatile (Robertson and Miller 1995a) it would concentrate along with the guanidinium. The Arrhenius curve shows that this synthesis would be effective at lower temperatures. At  $25^\circ C$ , the  $t_{1/2} = 50$  years with  $1\text{ }m$  guanidinium and 2.3 years with  $20\text{ }m$  guanidinium. The reaction would even work at  $0^\circ C$  ( $t_{1/2} = 113$  years at  $20\text{ }m$ ).

#### Drying Lagoons and Beaches as Prebiotic Reactors

The dried beach locale has been used by a number of workers studying prebiotic reactions (e.g., Fox and Dose 1972; Fuller et al. 1972; Keefe et al. 1995), and it is considered one of the best methods by which to concentrate dilute organic compounds in a lake or ocean (Miller and Orgel 1974). However, the course of such prebiotic evaporation has never been examined in detail. Freshwater lakes evaporate to dryness relatively cleanly but do contain some dissolved salts. The organic content of a prebiotic lake could not have been very high because of the short accumulation times. Thus a lake containing  $100\text{ }ppm$   $NaCl$  ( $1.7 \times 10^{-3}\text{ }M$ ) might have an organic/ $NaCl$  ratio not much different from the prebiotic ocean.

The evaporation process for seawater is more complex. When lagoons are isolated from the ocean or other water sources they dry out at a rate that depends on the temperature, relative humidity, and especially depth. The evaporation can be as much as several meters per year. The less soluble salts precipitate in succession, and very soluble compounds such as urea,  $G \cdot HCl$ , and cyanoacetaldehyde would precipitate near the end of the evaporation. The final concentration of  $G \cdot HCl$ , for example, would be the same  $23\text{ }m$  whether the initial concentration in the seawater is  $10^{-4}$  or  $10^{-8}\text{ }m$ , and the time for concentration to saturation is essentially the same because most of the increase in concentration occurs as the last portion of the water evaporates. Thus the initial concentration of  $G \cdot HCl$  is not particularly impor-

tant, although the amount of saturated  $G \cdot HCl$  solution would be correspondingly less. This may not hold for extremely low concentrations because adsorption processes may remove some of the  $G \cdot HCl$  before it can reach saturation.

The final stages of the evaporation to dryness are very complex. Complete drying out of seawater is difficult because the relative humidity of saturated  $MgCl_2$  and  $CaCl_2$  solutions is 33% (at  $25^\circ C$ ), and the organic solutes would lower the relative humidity even further. Thus the concentrated reactants (cyanoacetaldehyde and  $G \cdot HCl$ , in our case) would have considerable time to react before the solution completely dries out. Even with very low relative humidities the drying out of a lagoon would be very slow. The saturated solution between the precipitated salt crystals evaporates very slowly because it is not exposed to the atmosphere except at the top layer.

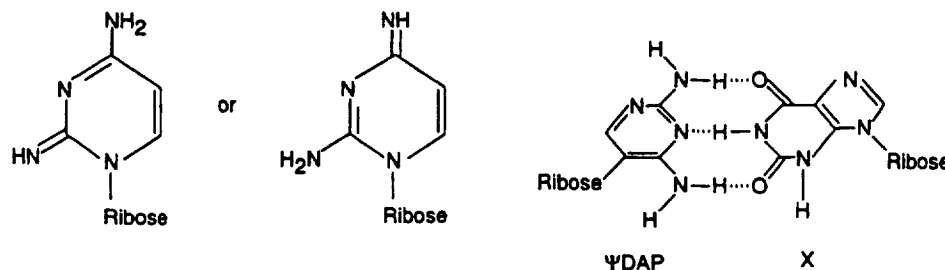
It is clear that the prebiotic chemistry of evaporating lagoons and beaches requires more investigation. However, our preliminary experiments show that short-term evaporation processes give yields of DAP that are high enough to be of prebiotic importance even though they are lower than the yields of sealed tube experiments.

#### DAP and SC as Sources of C and U

As shown previously (Ferris et al. 1974), DAP is not a good source of cytosine because the hydrolysis of cytosine to uracil is faster than the hydrolysis of DAP to C. DAP is an even poorer source of IC, the ratio of IC/C being about 0.25. It is possible that the pH-rate profile and temperature coefficients of the four rate constants are such that cytosine is more abundant at lower temperatures than at  $100^\circ C$ .

In the efficient prebiotic synthesis of cytosine from cyanoacetaldehyde and urea, the hydrolytic instability of cytosine is also a problem (Shapiro and Klein 1966; Garrett and Tsau 1972). If the first genetic material used only adenine and uracil, then either cytosine or DAP would be an effective precursor. But this two-letter code has a similar problem with the adenine: It hydrolyzes to hypoxanthine at a rate  $1/20$  that of cytosine (Shapiro 1995). Thus the adenine of the AU base pair needs to be used within a few thousand years after their synthesis if the temperature is  $-25^\circ C$ .

Diaminopyrimidine cannot be used as a riboside because the ribose forces the exocyclic amines into the imino forms.



However, a ribose or a precursor backbone on the 5-carbon ( $\Psi$ DAP) could form a 3-hydrogen-bond base pair with xanthine (X) as proposed by Piccirilli et al. (1990). Since DAP is more stable to hydrolysis than C and not much faster to hydrolyze than adenine, a  $\Psi$ DAP-xanthine base pair has some advantages over uracil-adenine. There would be three hydrogen bonds as there is with the guanine-cytosine base pair.

Isocytosine may have played a prebiotic role since it forms three hydrogen bonds with isoguanine (Switzer et al. 1989; Piccirilli et al. 1990), even though its steady-state concentration is lower than cytosine. The kinetics of the decomposition of isoguanine would be needed to evaluate this possibility.

We have been unable to detect the hydroxymethylation of DAP with formaldehyde, a reaction that is efficient with uracil and works well but more slowly with cytosine, uridine, and cytidine (Robertson and Miller 1995b). It is possible that the Mannich reaction would work (e.g. the addition of  $\text{CH}_2=\text{N}-\text{CH}_2\text{COO}^-$  to the C-5 of DAP) thereby producing a nonribose precursor to  $\Psi$ DAP.

Although we obtained SC from cyanoacetaldehyde and thiourea, the yields were quite low, a factor of 100 less than with urea or  $\text{G} \cdot \text{HCl}$ . Nor did SC form from  $\text{H}_2\text{S}$  and DAP. There may be other ways to get 2-thiopyrimidines that are more efficient, but if not, then thiopyrimidines may still have been important bases in the pre-RNA world if their use was based on factors other than abundance.

The lower yields might be overcome if thiourea were more abundant than urea or G, but this is not likely except in special areas. SC is a better source of SU than C, although both hydrolyze relatively rapidly to U. Since C is produced directly from cyanoacetaldehyde and urea and to a lesser extent from DAP, C is likely to have been more abundant than SU, but the C/SU may not have been that high. The end product of these hydrolyses, U, is likely to have been the most abundant pyrimidine.

Although SC is not likely to have been very abundant in the preRNA world, SU and 5-substituted SUs may have been important. Thiouridines stabilize double-stranded RNA and form special structures in the anticodon loop of tRNAs. There may have been other roles in the RNA world. Thus the versatility of RNA in the RNA world would have been considerably expanded by the availability of thiouracils.

*Acknowledgments.* We thank Jason Dworkin, Charles Perrin, and Anthony Keefe for helpful comments and the NASA Specialized Center for Research and Training in Exobiology (NSCORT) for grant support.

## References

- Elvers B et al. (eds) (1989) Ullmann's encyclopedia of industrial chemistry, vol A12. 5th ed. VHC, New York, p 546
- Fabbrizzi L, Micheloni M, Paoletti P, Schwarzenbach G (1977) Protonation process of unusual exothermicity. *J Am Chem Soc* 99: 5574-5576
- Ferris JP, Sanchez RA, Orgel LE (1968) Studies in prebiotic synthesis III. Synthesis of pyrimidines from cyanoacetylene and cyanate. *J Mol Biol* 33:693-704
- Ferris JP, Goldstein G, Beaulieu DJ (1970) Chemical evolution IV. An evaluation of cyanovinyl phosphate as a prebiotic phosphorylating agent. *J Am Chem Soc* 92:6598-6603
- Ferris JP, Zamek OS, Altbuch AM, Freiman H (1974) Chemical evolution XVIII. Synthesis of pyrimidines from guanidine and cyanoacetaldehyde. *J Mol Evol* 3:301-309
- Fuller WD, Sanchez RA, Orgel LE (1972) Studies in prebiotic synthesis VII. Solid-state synthesis of purine nucleosides. *J Mol Evol* 1:249-257
- Fox SW, Dose K (1972) Molecular evolution and the origin of life. WH Freeman, San Francisco
- Frost AA, Pearson RG (1961) Kinetics and mechanism. John Wiley, New York, pp 166-169
- Garrett ER, Tsau J (1972) Solvolyses of cytosine and cytidine. *J Pharm Sci* 61:1052-1061
- Keefe AD, Newton GL, Miller SL (1995) A possible prebiotic synthesis of pantetheine, a precursor to coenzyme A. *Nature* 373:683-685
- Kirk RE, Othmer DF (1951) Encyclopedia of chemical technology, vol. 7. Interscience Encyclopedia, New York, p 326
- Limbach PA, Crain PF, McCloskey JA (1994) Summary: the modified nucleosides of RNA. *Nucleic Acids Res* 22:2183-2196
- Lohrmann R (1972) Formation of urea and guanidine by irradiation of ammonium cyanide. *J Mol Evol* 1:263-269
- Lowe CU, Rees MW, Markham R (1963) Synthesis of complex organic compounds from simple precursors: formation of amino-acids, amino-acid polymers, fatty acids and purines from ammonium cyanide. *Nature* 199:219-222
- Miller SL (1957) The mechanism of synthesis of amino acids by electric discharges. *Biochim Biophys Acta* 23:480-489
- Miller SL, Orgel LE (1974) The origins of life on the earth. Prentice-Hall, Englewood Cliffs, NJ, pp 129-134
- Piccirilli JA, Krauch T, Moroney SE, Benner SA (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* 343:33-37
- Reid EE (1963) Organic chemistry of bivalent sulfur. vol. 5. Chemical Publishing Co. Inc., New York, pp 11-17
- Robertson MP, Miller SL (1995a) An efficient prebiotic synthesis of cytosine and uracil. *Nature* 375:772-774
- Robertson MP, Miller SL (1995b) Prebiotic synthesis of 5-substituted uracils: a bridge between the RNA world and the DNA-protein world. *Science* 268:702-705
- Sanchez RA, Ferris JP, Orgel LE (1966) Cyanoacetylene in prebiotic synthesis. *Science* 154:784-785
- Shapiro R (1995) The prebiotic role of adenine: a critical analysis. *Orig Life Evol Biosph* 25:83-98
- Shapiro R, Klein RS (1966) The deamination of cytidine and cytosine by acidic buffer solutions. Mutagenic implications. *Biochemistry* 5:2358-2362
- Shaw WHR, Walker DG (1958) Kinetic studies of thiourea derivatives. IV. The methylated thioureas. Conclusions. *J Am Chem Soc* 80: 5337-5342
- Shnidman L (1933) The solubility of thiourea in water, methanol, and ethanol. *J Phys Chem* 37:693-700
- Switzer C, Moroney SE, Benner SA (1989) Enzymatic incorporation of a new base pair into DNA and RNA. *J Am Chem Soc* 111:8322-8323