Title: Towards Self-Replicating Chemical Systems Based on Cytidylic and Guanylic Acids

Final Report for the Period October 15, 1997 to January 14, 1999

(including the extension from Jan 15 to Jul 14, 1999)

This project was aimed towards a better understanding of template-directed reactions and, based on this, towards the development of efficient non-enzymatic RNA replicating systems. These systems could serve as models for the prebiotic synthesis of an RNA world. The major objectives of this project are: (a) To elucidate the mechanistic aspects of template-directed (TD) chemistry and (b) to identify active boundary regions, or conditions, environmental and other, that favor "organized chemistry" and stereoselective polymerization of nucleotides. "Organized chemistry" may lead to enhanced polymerization efficiency which in turn is expected to facilitate the road towards a self-replicating chemical system based on all four nucleic acid bases.

During the first year of the granting period we completed objective (b) and during the second year and the extension given, specifically during October 1997 to July 1999, we completed objective (a). The research resulted in the publication of three papers, # 1 to 3 (# 1, identified as submitted and # 2 identified as manuscript in preparation in the previous report) and two papers are in press, # 4 and 5. Reprints of # 1 to 3 are enclosed and the abstracts of # 4 and 5 are included in the Appendix of this report.

List of publications and manuscripts.

1. A. Kanavarioti, "Preference for Internucleotide Linkages as a Function of the Number of Constituents in a Mixture," *J. Mol. Evol.*, **46**, 622-632 (1998).
2. A. Kanavarioti, C. F. Bernasconi, E. E. Baird, "Effects of Monomer and Template Concentration on the Kinetics of Nonenzymatic Template-Directed Oligoguanylate Synthesis," *J. Am. Chem. Soc.*, **120**, 8575-8581 (1998).
3. A. Kanavarioti, "Kinetic Preference for the 3'-5'-Linked Dimer in the Reaction of Guanosine 5'-phosphoryl-morpholinamide with Deoxyguanosine 5'-phosphoryl-2-methylimidazolide as a Function of Poly(C) Concentration", *J. Org. Chem.*, **63**, 6830-6838 (1998).
4. A. Kanavarioti, L. F. Lee, S. Gangopadhyay, "Relative Reactivity of Ribosyl 2'-OH vs. 3'-OH in Concentrated Solutions of Phosphoimidazolide Activated Nucleotides. *Origins Life Evol. Biosph.*, in press.
5. A. Kanavarioti, E. E. Baird, T. B. Hurley, J. A. Carruthers, S. Gangopadhyay, "Unique Catalysis and Regioselectivity Observed in the Poly(C)-directed RNA-dimer Formation from 2-MeImpG. Kinetic Analysis as a Function of Monomer and Polymer Concentration. *J. Org. Chem.*, in press.

Participation in meetings:

Participated in the 6th Symposium on Chemical Evolution and the Origin and Evolution of Life held in November 17-20, 1997 at NASA/Ames Research Center, CA.

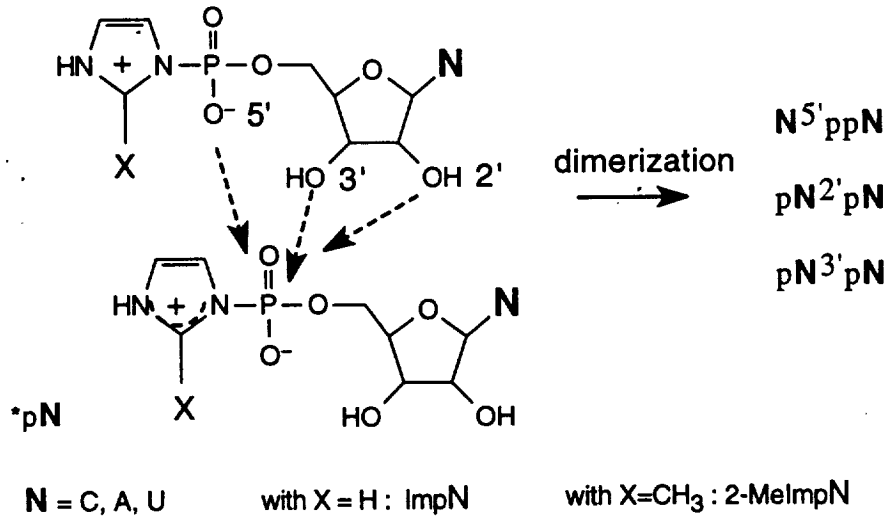
Attended the 27th Reaction Mechanisms Conference held June 28 to July 3, 1998 in Asilomar, Pacific Grove, CA.

Presented a paper at the 14th International Conference on Physical Organic Chemistry held on August 16-21, 1998, in Florianopolis, SC, Brazil.

Presented two papers at the 12th International Conference on the Origin of Life and 9th ISSOL Meeting held on July 11-16, 1999 in San Diego, CA.

SUMMARY OF THE RESEARCH ACTIVITIES DURING THE ABOVE PERIOD:

In the first year of this cooperative agreement we synthesized a number of activated nucleotides (abbreviated *pN, see structure below) with different leaving groups and different nucleobases. We also developed techniques for analysis of mixtures by high performance liquid chromatography (HPLC) and a combination of liquid chromatography/mass spectrometry (LC/MS). During the period identified in this report, we performed a large number of experiments with the synthesized substrates. Kinetic analysis of the experiments unraveled some fascinating aspects of nucleotide chemistry that will be elaborated upon below. The research was conducted by Dr. Kanavarioti with the assistance of an undergraduate student, Lynn F. Lee, and by consultation with Prof. C. F. Bernasconi of UCSC and Dr. S. Chang of NASA/Ames. In the 2nd part of the above identified period our Ames consultant was Dr. Mark Fonda who took over when Dr. S. Chang retired. It should be mentioned that in June of 1998 Dr. Sumana Gangopadhyay, a Visiting Postdoctoral Researcher joined the research group who worked on a part-time basis. She is on leave of absence from the Department of Chemistry, Gurudas College, Calcutta, India.



N = C, A, U with X = H : ImpN with X=CH_3 : 2-MeImpN

Scheme 1: Structure and Reactivity of Phosphoimidazolide Activated Nucleotides, abbreviated *pN.

The Evaporating Lagoon Scenario with Mononucleotides.

It is one of our theses that optimization of the conditions that lead to stacking of monomers in aqueous solutions will facilitate optimization of conditions for TD chemistry. Oligomerization yields suffer because of competition by hydrolysis. Although not generally agreed upon, solutions with high concentration of organic molecules are plausible in an evaporating lagoon or pools on drying beaches on the early Earth.[1,2] High yields of cytosine were obtained from the reaction of cyanoacetaldehyde in concentrated urea solutions.[1] In the presence of metal ions self-condensation of *pC and *pU derivatives (Scheme 1) yields about 30 % of condensation products at 0.1 M and about 60 % at 1.0 M monomer concentration.[2] Products consist of dimers and decreasing amounts

of oligomers up to the 6-mer. The yield of internucleotide-linked dimers and longer oligomers increase with monomer concentration. Surprisingly, the yield of pyrophosphate dimer, typically the dimer formed the easiest,[3] does not increase.[2] The high yield of condensation products formed in these concentrated solutions makes the evaporating lagoon scenario a potentially interesting 'active boundary region' for prebiotic synthesis of short oligomers.

Weak Cooperativity in Pyrimidine Self-stacking.

Stacking of nucleotides in water provides the major force for self-asscociation as shown by H^1 NMR upfield shifts of the ring protons, decreasing molar osmotic coefficients and hypochromic effects in the UV as a function of increasing concentration of the nucleotide.[4] The relatively larger area of the purine bases compared to the pyrimidine bases leads to stronger complexation by stacking in the order purine•purine > purine•pyrimidine > pyrimidine•pyrimidine. Available data using different methods fit best an isodesmic model where the association constant of one molecule, such as cytidine (C), to another, K_1, is equal to the association constant, K_{n-1}, of the n-th cytidine monomer to a stack of n-1 cytidine units where $n \geq 2$ ($K_{n-1}=K \approx 0.5\ M^{-1}$ for 5'CMP and $K_{n-1}=K \approx 0.25\ M^{-1}$ for 5'UMP).[5] It is only with guanosine derivatives where self-association is shown to fit a cooperative model where $K_1<K_2=K_n$.[4]

$$C + C \overset{K_1}{\Leftrightarrow} C_2,\quad C_2 + C \overset{K_2}{\Leftrightarrow} C_3,\quad C_3 + C \overset{K_3}{\Leftrightarrow} C_4, \ldots\ldots\quad C_{n-1} + C \overset{K_{n-1}}{\Leftrightarrow} C_n$$

In exploring the chemistry of *pN we observed that the percent yield of dimers and short oligomers increases as a function of initial monomer concentration up to 0.3 to 0.4 M. However, at concentration of monomer higher than 0.4 M the product distribution becomes practically independent of substrate concentration even at early reaction times. The observation of first-order kinetics for product formation can be rationalized by self-association that results in practically complete incorporation of the substrate in complexes.

We have used the yield of condensation product as a measure of the fraction of the substrate being incorporated in complexes or stacks. This was done by expressing the condensation yield in monomer equivalents and accounting for the fact that oligomerization is in competition with hydrolysis. Analysis of product yield as a function of substrate concentration allows calculation of self-association constants for adenosine, cytidine and uridine derivatives. Computer simulations were performed on three basic models: The isodesmic model with $K_1=K_2=K_{n-1}$, a model of self-association which leads to a two-unit complexation, with $K_1 \neq 0$ and $K_2=K_3=K_{n-1}=0$, and the cooperative model with $K_1<K_2=K_3=K_{n-1}$. Attempts were made with another cooperative model where $K_1<K_2<K_3=K_4=K_{n-1}$ but this did not improve the fit. Under most conditions tested the data with adenosine, cytidine and uridine derivatives fit the coooperative model and allow an estimate of K_1 and K_{n-1}. For example, we deduce $K_1=2\ M^{-1}$ and $K_{n-1}=6\ M^{-1}$ with 2-MeImpU at pH 7.45, 20° C and in the presence of 0.2 M Mg^{2+}. Not only the affinity for stacking is surprisingly high (see literature values above) but the interpretation regarding cooperativity is unprecedented for the two pyrimidines. The enhanced self-association and the switch from non-

cooperative to cooperative complexation observed with *pN, compared to the parent compounds, is tentatively attributed to the presence of the imidazole moiety exerting additional intermolecular stacking interactions in the form of imidazole/nucleobase or imidazole/imidazole. Corraborative evidence for this effect comes from molecular dynamics (MD) simulations with monomers (see proposed Work for JRI NCC 2-5309).

Selective Amplification in the Formation of Internucleotide Linkages: U,C,G Mixtures.

We have studied product distributions in reactions with one, two or three reactive components at the same total nucleotide concentration.[6] *pN used as substrates were the nucleoside 5' phosphate 2-methylimidazolides, 2-MeImpN with N = cytidine (C), uridine (U) or guanosine (G). Reactions were conducted as self-condensations, i.e. one nucleotide only; with two components in the three binary U,C-, U,G- and C,G-mixtures and with three components in the ternary U,C,G-mixture. The products are 5'NMP, 5',5'-pyrophosphate-, 2',5'-, 3',5'-linked-dimers (see Scheme 1), cyclic dimers and a small percentage of longer oligomers. The surprising finding was that, under identical conditions, including the same total monomer concentration, the product distribution differs substantially from one reaction to another, most likely due to changing intermolecular interactions depending on the constituents. Even more unexpected was the observed trend according to which reactions of the U,C,G-mixture produce the highest yield of internucleotide-linked dimers, whereas the self-condensations produce the least and the reactions with the binary mixtures produce yields that fall in between. What is remarkable is that the approximately two-fold increase of the percent yield of internucleotide-linked dimers is not due to a concentration effect or a catalyst, but it is due to the increased complexity of the system from a single to two and three components.[6] These observations indicate selective amplification of internucleotide linked products with increased complexity in relatively simple chemical systems.

The chemistry of *pN has unraveled a number of interesting avenues such as enhanced yields of internucleotide-linked products, cooperative stacking exhibited by all four nucleotides and selective amplification of internucleotide-linked products with increased number of components in a mixture. We wish to explore these phenomena further in order to develop models of active boundary regions in which chemical evolution could have occurred and because of their relevance for monomer incorporation in TD synthesis.

REFERENCES:

1. Robertson, M. P. and Miller S. L. An efficient prebiotic synthesis of cytosine and uracil. *Nature* **1995,** *375*, 772-774, and correction: *377*, 257.
2. Kanavarioti, A. Dimerization in highly concentrated solutions of phospho-imidazolide activated mononucleotides. *Origins Life Evol Biosph* **1997,** *27*, 357-376.
3. Rodriguez, L. & Orgel L. E. Pyrophosphate formation as the most efficient condensation reaction of activated nucleotides. *J. Mol. Evol.*, **1991,** *32*, 101-104.
4. Saenger, W. **(1984)** Principles of Nucleic Acid Structure; Springer-Verlag: New York, Inc. for syn-, anti-conformation see p 69, for hydrogen-bonding see p. 120, for stacking interactions see p 134.
5. Tribolet, R.; Sigel, H. Self-association of adenosine 5'-monophosphate (5'-AMP) as a function of pH and in comparison with adenosine, 2'-AMP and 3'-AMP. *Biophys. Chem.* **1987,**

27, 119-130. Walmsley, J. A.; Sagan, B. L. The Effect of Monovalent Cations on the Association Behavior of Guanosine 5'-Monophosphate, Cytidine 5'-Monophosphate, and Their Equimolar Mixture in Aqueous Solution. *Biopolymers*, **1986**, *25*, 2149-2172. Mitchell, P. R.; Sigel, H. A Proton Nuclear-Magnetic-Resonance Study of Self-Stacking in Purine and Pyrimidine Nucleosides and Nucleotides. *Eur. J. Biochem.* **1978**, *88*, 149-154. Ts'o, P. O. P. ed. **1974.** *Basic Principles in Nucleic Acid Chemistry*. 2 vols. New York: Academic Press (chapter 6 of vol 1). Martin, R. B. Comparisons of Indefinite Self-Association Models. *Chem. Rev.* **1996**, *96*, 3043-3064. Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.-F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C. and Kool, E. T. Experimental measurement of aromatic stacking affinities in the context of duplex DNA. *J. Amer. Chem. Soc.* **1996,** *118,* 8182-83.

6. Kanavarioti, A. Preference for Internucleotide Linkages as a Function of the Number of Constituents in a Mixture, *J. Mol. Evol.*, **1998,** *46*, 622-632.

Highlight: The most important finding of this work is **evidence for weak cooperativity in the self-association of phosphoimidazolide-activated pyrimidine nucleotides in water and in the presence of divalent metal ions.**

In conclusion, we believe that we completed satisfactorily the objectives set in this cooperative agreement and in view of the recent retirement of Dr. Sherwood Chang we requested the closure of this agreement at the end of the 2nd year. We did request and obtained a six-month extension in order to complete and submit manuscripts #4 and 5. The results are described in the publications, submitted manuscripts and in this report. They should enhance our understanding of the mechanisms that may have operated and led to chemical evolution on early Earth and on other similar extraterrestrial bodies. At this time leads coming of the above work are being pursued as part of a JRI with Dr. Andrew Pohorille of NASA/Ames (NCC 2-5309 initiated on Feb 1, 1999 entitled "From Prebiotic Chemistry to an RNA World").

Appendix

 # 4 Abstract : Phosphoimidazolide activated ribomononucleotides (*pN, see structure) are useful substrates for the non-enzymatic synthesis of oligonucleotides. In the presence of metal ions, aqueous solutions of *pN typically yield dimers and small amounts of higher oligomers. Prominent among the condensation products are the two internucleotide linked dimers: $pN^{2'}pN$ and $pN^{3'}pN$. Here the relative nucleophilicity of 2′-OH vs. 3′-OH was determined from the ratio of the yields of $pN^{2'}pN$ vs. $pN^{3'}pN$. Experiments were performed at 23° C in the range $7.2 \leq pH \leq 8.4$ with substrates that differ in nucleobase (guanosine (G), cytidine (C), uridine (U), and adenosine (A)) and leaving group (imidazole (Im), 2-methylimidazole (2-MeIm) and 2,4-dimethylimidazole (2,4-diMeIm)). In addition, two metal ions (Mg^{2+} or Mn^{2+}) were employed as catalysts. Other variables were the substrate concentration (0.1 M to 1.0 M), and the metal ion concentration (0.05 M to 0.2 M). Typically, the product ratio $pN^{2'}pN$: $pN^{3'}pN$ = 2′-5′: 3′-5′ varied between 2 to 3 indicating that the 2′-OH is about 2 to 3 times more nucleophilic than the 3′-OH. This conclusion contradicts earlier studies reporting a relative nucleophilicity 2′-5′: 3′-5′= 6 to 9 obtained under somewhat different conditions. Possible reasons for this discrepancy are discussed. In addition, tips are offered for a high purity, one-pot, synthesis of *pN.

 # 5 Abstract : Polycytidylate, poly(C), serves as a scaffold or template to direct and catalyze the synthesis of long oligoguanylates from guanosine 5'-phosphate 2-methylimidazolide, 2-MeImpG. In the absence of poly(C) small amounts of three isomeric dimers, i.e. the 2'-5'-, the 3'-5'-, and the pyrophosphate-linked, are formed slowly. In the presence of poly(C) oligomers that are primarily 3'-5'-linked are formed quickly and in high yield. Product analysis suggests that the oligomers are elongation products of the 3'-5'-linked dimer, abbreviated D. Assuming that D is formed slowly from two molecules of 2-MeImpG (Scheme 1) and elongates relatively fast, the initial rate of dimerization, d[D]/dt in Mh^{-1}, was determined using two independent methods. The first method is based on the approximation that at the onset of the reaction the substrate is consumed only via hydrolysis and dimerization and thus elongation can be neglected. The second, more accurate, method exploits the assertion that every oligomer was once a 3'-5'-linked dimer. Hence the concentration of D was obtained indirectly from the concentration of the oligomer products. These two methods gave comparable results. Experiments were run in aqueous solution in the presence of 1.0 M NaCl, 0.2 M $MgCl_2$ at pH 7.9±0.1 and 23° C. Controls were run in the absence of poly(C) and in the presence of other polynucleotides. The kinetics were determined as a function of both monomer and polymer concentration the latter expressed in C equivalents.

The kinetic data obtained in the presence of poly(C) confirmed an earlier conclusion regarding the remarkable effect of poly(C) on the formation of the 3'-5'-linked diguanylate. Initial dimerization rates were quantitatively correlated using a simple template-directed (TD) model that presumes cooperative binding (two association constants) of 2-MeImpG on poly(C) and reaction between adjacent template-bound molecules. The model allows for the estimation of the asscociation constants and the intrinsic rate constant of dimerization, k_2*. Insights into the detailed mechanism are also gained from this analysis. The fact that the proposed model can successfully

UNIVERSITY OF CALIFORNIA, SANTA CRUZ

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

SANTA CRUZ, CALIFORNIA 95064

August 18, 1999

Ms. Beatrice Morales
Grants Officer
Space and University Affairs
Mail stop: 241-1
NASA/Ames Research Center
Moffett Field, CA 94035-1000

Re: Final Report for Grant No. NCC 2-534

Enclosed you will find the final report for the cooperative agreement NCC 2-534 entitled "Towards Self-Replicating Chemical Systems Based on Cytidylic and Guanylic Acids" covering the period October 15, 1997 to January 14, 1999 including a six-month no-cost extension to July 14, 1999. No patents were submitted based on the research done and no equipment has been borrowed from Ames Research Center under this agreement. If you need additional information please contact me or Karen Fry.

Sincerely yours,

A Kanavarioti

Dr. Anastassia Kanavarioti, P.I.
Associate Research Chemist
(408) 459-2978
e-mail: tessi@hydrogen.ucsc.edu

cc: (i) Dr. Mark Fonda, NASA/Ames
(ii) CASI
(iii) Karen Fry, Contracts and Grants Officer, UCSC

NCD. AUG 23 1999

CC: 202A-3
203-18
255-2

CAST