Research Paper

Sequence Analysis of Trimer Isomers Formed by Montmorillonite Catalysis in the Reaction of Binary Monomer Mixtures

GÖZEN ERTEM,1,2,3 ROBERT M. HAZEN,1 and JASON P. DWORKIN2

ABSTRACT

Oligonucleotides are structurally similar to short RNA strands. Therefore, their formation via non-enzymatic reactions is highly relevant to Gilbert’s RNA world scenario (1986) and the origin of life. In laboratory synthesis of oligonucleotides from monomers, it is necessary to remove the water molecules from the reaction medium to shift the equilibrium in favor of oligonucleotide formation, which would have been impossible for reactions that took place in dilute solutions on the early Earth. Model studies designed to address this problem demonstrate that montmorillonite, a phyllosilicate common on Earth and identified on Mars, efficiently catalyzes phosphodiester-bond formation between activated mononucleotides in dilute solutions and produces RNA-like oligomers. The purpose of this study was to examine the sequences and regiospecificity of trimer isomers formed in the reaction of 5'-phosphorimidazolides of adenosine and uridine. Results demonstrated that regiospecificity and sequence specificity observed in the dimer fractions are conserved in their elongation products. With regard to regiospecificity, 61% of the linkages were found to be RNA-like 3',5'-phosphodiester bonds. With regard to sequence specificity, we found that 88% of the linear trimers were hetero-isomers with 61% A-monomer and 39% U-monomer incorporation. These results lend support to Bernal’s hypothesis that minerals may have played a significant role in the chemical processes that led to the origin of life by catalyzing the formation of phosphodiester bonds in RNA-like oligomers. Key Words: Origin of life—RNA—HPLC—Mineral catalysis—Montmorillonite—Oligonucleotides. Astrobiology 7, 715–722.

INTRODUCTION

THE POSSIBLE ROLE OF MINERALS as a catalyst in the processes that lead to the abiotic formation of bio-molecules from monomers was first proposed by Bernal as early as 1949 (Bernal, 1949). Research carried out to test this hypothesis, namely, the role of mineral catalysis on the formation of RNA-like oligomers, has been reviewed (Ertem, 2004; Ferris, 2006; Shapiro, 2006). Model studies demonstrate that the clay mineral montmorillonite catalyzes the condensation of activated monomers in an aqueous electrolyte solution at pH 8, which produces RNA-like

1Geophysical Laboratory and NASA Astrobiology Institute, Washington, DC.
2Astrochemistry Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland.
3Space Science Division, NASA Ames Research Center, Moffett Field, California.
oligomers* with 10–12 monomer units (Ferris and Ertem, 1992a; Ding et al., 1996; Ertem and Ferris, 1996, 2000) and up to 50 monomer units in length (Ferris et al., 1996; Ferris, 2002), depending upon reaction conditions and type of activated monomer used. It has also been shown that oligocytidylates formed by montmorillonite catalysis serve as templates for the formation of complementary oligoguanylates (Ertem and Ferris, 1996). This is the first account demonstrating the template capability of a synthetic oligonucleotide chain for producing the complementary strands.

Self-condensation of activated monomers in concentrated reaction media associated with ice matrices produces oligouridylates up to 11 units long (Kanavarioti et al., 2001).

Montmorillonite, which was used as a catalyst in our studies, is one of the most abundant clay minerals on Earth. It is a member of the phyllosilicate group minerals, which were identified on Mars by OMEGA, a visible–near infrared hyperspectral imager (Bibring et al., 2005). Montmorillonite has a dioctahedral layer structure, in which each layer is made up of one octahedral sheet sandwiched between two tetrahedral sheets (Grim, 1968). The structural formula of the homoionic sodium form of the clay (Na-montmorillonite), derived from its elemental analysis (Brindley and Ertem, 1971) following the Ross and Hendricks model (1945), illustrates that main substitutions occur in the octahedral sheet, as shown below, where [ ]IV and [ ]VI represent tetrahedrally and octahedrally coordinated cations, respectively:

$$[\text{Si}_{3.89} \text{Al}_{0.11}]^{\text{IV}}[\text{Al}_{1.57} \text{Fe}_{0.173}^{3+} \text{Fe}_{0.022}^{2+} \text{Mg}_{0.27}]^{\text{VI}}\text{O}_{10}(\text{OH})_2$$

**Tetrahedral sheet**

**Octahedral sheet**

---

*We use the term oligomer to refer to any product that is produced in the condensation of mononucleotides by the formation of a phosphodiester bond and can be as short as two monomer units long.

---

FIG. 1. (a) A short segment of RNA strand and conventional shorthand notation showing its sequence. (b) A short segment of a synthetic RNA-like oligonucleotide that formed in the reaction of activated monomers by montmorillonite catalysis. N represents any nucleobase, in this case, adenine (A) or uracil (U).
Montmorillonite owes its catalytic activity to the isomorphic substitutions, i.e., substitution of Al\(^{3+}\) ions by Fe\(^{2+}\) and Mg\(^{2+}\) ions in the octahedral sheet. The negative charge excess thus created is counterbalanced by the interlayer cations, such as Ca\(^{2+}\) ions and Na\(^{+}\) ions, which are held between the layers and are easily exchangeable with other cations and organic molecules. Detailed studies have established, oligomerization reactions, which we discuss in this work, mainly take place on the basal surfaces of montmorillonite (Ertem and Ferris, 1998).

Figures 1a and 1b illustrate, respectively, the structures of a short strand of RNA and an RNA-like synthetic oligonucleotide segment formed in a montmorillonite-catalyzed condensation of activated mononucleotides, i.e., 5'-phosphorimidazolide of nucleosides (ImpN). The main difference between the two structures is that while contemporary RNA contains, with few exceptions, only 3',5'-linkages, monomer units in synthetic oligomer strands are linked together by 3',5'- as well as 2',5'-phosphodiester bonds.

Sequence analysis of dimer fractions formed in a montmorillonite-catalyzed reaction of quaternary activated monomer mixtures (ImpNs) shows a common trend regarding the regioselectivity and sequence selectivity of phosphodiester bonds, and a large variation in their yields (Ertem and Ferris, 2000). In the binary reaction of A- and U-monomers, 73% of the isomers are heterodimers, i.e., contained both A and U monomers, and 51% of the bonds joining the monomer units together are RNA-like 3',5'-linkages. The objective of this research is to determine whether the heterogeneity and regioselectivity observed in dimer isomers are preserved in the trimer and higher isomers.

**EXPERIMENTAL SECTION**

Chemicals and HPLC-grade solvents were purchased from the same sources listed in a previous work (Ferris and Ertem, 1993). Imidazole and Polyadenylic, uridylic acid (Poly[A,U]) were from Sigma, alkaline phosphatase enzyme (APH) was from Worthington. Volclay SPV-200 was a gift from American Colloid Co., Arlington Heights, Illinois.

Activated mononucleotides were synthesized following the procedure described by Joyce et al. (1984), with slight modifications (Ferris and Ertem, 1992b). Volclay was converted into its homoionic Na-form by the titration method (Banin et al., 1985).

High Performance Liquid Chromatography (HPLC) analysis was carried out on an Agilent 1050 instrument. Biosphere GMB 1000 Q anion exchange column was obtained from Puresyn Inc., Malvern, Pennsylvania. For reverse phase HPLC, Alltech’s C-18 Alltima column was used. HPLC conditions were the same as described before (Ertem and Ferris, 2000).

**Preparation of standards**

The homo-trimers N\(^2\)pN\(^2\)pN, N\(^3\)pN\(^2\)pN, N\(^3\)pN\(^3\)pN were isolated from the products of montmorillonite-catalyzed self-condensation reactions of ImpA and ImpU.

Hetero-trimers A\(^3\)pA\(^3\)pU, A\(^3\)pU\(^3\)pU, U\(^3\)pU\(^3\)pA, U\(^3\)pA\(^3\)pA, A\(^3\)pU\(^3\)pA, and U\(^3\)pA\(^3\)pU were isolated from the partial base hydrolysis products of Poly[A,U]: 1 mg of Poly[A,U] was dissolved in 1 mL of 0.1 M NaOH and hydrolyzed for 40–50 min at 70°C to produce (N\(^3\))\(^p\) type oligomers (Lohrmann et al., 1980). Since the distribution of A- and U-monomers in Poly[A,U] strands is random, the trimer fraction separated from its partial hydrolysis products using an anion exchange column contained all six of the hetero-isomers along with A\(^3\)pA\(^3\)pA\(^3\)p and U\(^3\)pU\(^3\)pU\(^3\)p. The trimer fraction, which contained N\(^3\)pN\(^3\)pN\(^3\)p isomers, where N is A or U, was collected from the anion exchange column and digested with APH to hydrolyze the 3'-phosphate groupings. The N\(^3\)pN\(^3\)pN mixture thus formed was dialyzed, freeze-dried, dissolved in water, and injected on a reverse phase column. As expected, eight isomers appeared on the chromatogram. To identify their sequence, each isomer was collected separately and hydrolyzed with Ribonuclease T\(_2\) (RNase T\(_2\)). Hydrolysis products were identified by co-injections with authentic standards using reverse phase HPLC. RNase T\(_2\) is an enzyme that hydrolyzes 3',5'-linkages at pH 4–5 by first forming a 2',3'-cyclic isomer followed by a hydrolytic cleavage that leads to the formation of 3'-phosphate derivative. RNase T\(_2\) hydrolysis products were digested with APH to cleave the 3'-phosphate groupings, and products thus formed were co-injected with authentic standards for identification. Table 1 lists the hydrolysis products of each isomer.
toward cyclic isomers and pyrophosphate derivatives, as shown below:

\[
[pNpNpN]_4^{1^-} + [c-(pN)4]_4^{1^-} + [NpNppNpN]_4^{1^-} \\
\longrightarrow
\text{linear isomers} \quad \text{cyclic isomers} \quad \text{pyrophosphate derivatives}
\]

\[
[NpNpN]_2^{2^-} + [c-(pN)_3]_4^{2^-} + [NpNppNpN]_4^{1^-} \\
\text{hydrolyzed cyclic isomers} \quad \text{pyrophosphate derivatives}
\]

Linear trimers with the general formula of \(NpNpN\), where \(N = A\) or \(U\) for \(\text{ImpA-ImpU}\) reaction, were collected and further separated by reverse phase HPLC (Fig. 2).

### Fraction collection and enzymatic hydrolysis

For enzymatic hydrolysis studies, fractions were collected, either individually or two together—depending upon the difference between their close retention times, fractions 4 and 5; 6 and 7 and 8 were collected together. They produced different hydrolysis products.

### TABLE 1. THE PRODUCTS FORMED FROM RNase T2 HYDROLYSIS FOLLOWED BY APH HYDROLYSIS OF EACH TRIMER FRACTION OBTAINED FROM THE PARTIAL HYDROLYSIS OF POLY[A,U]

<table>
<thead>
<tr>
<th>Starting isomer</th>
<th>RNase T2 hydrolysis product</th>
<th>APH hydrolysis product</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A^3pA^3pA)</td>
<td>(A^3p + A^3p + A)</td>
<td>(A + A + A)</td>
</tr>
<tr>
<td>(A^3pA^3pU)</td>
<td>(A^3p + A^3p + U)</td>
<td>(A + A + U)</td>
</tr>
<tr>
<td>(A^3pU^3pU)</td>
<td>(A^3p + U^3p + U)</td>
<td>(A + U + U)</td>
</tr>
<tr>
<td>(U^3pA^3pA)</td>
<td>(U^3p + U^3p + A)</td>
<td>(U + U + A)</td>
</tr>
<tr>
<td>(U^3pA^3pU)</td>
<td>(U^3p + A^3p + U)</td>
<td>(U + A + U)</td>
</tr>
<tr>
<td>(U^3pU^3pU)</td>
<td>(U^3p + U^3p + U)</td>
<td>(U + U + U)</td>
</tr>
</tbody>
</table>

These data were used to establish the elution order of isomers from the reverse phase column. The elution order of 2’-linked analogs of these trimers, which are not commercially available, has been determined to be the same but with shorter retention times (Ertem and Ferris, 2000).

### Reactions

Reaction mixtures were prepared to contain each activated monomer at a final concentration of 0.014 \(M\) in 0.1 \(M\) HEPES, 0.2 \(M\) NaCl, 0.075 \(M\) MgCl\(_2\), pH 8, and Na-montmorillonite, according to Ertem and Ferris (2000). At the end of seven days, supernatants were separated by centrifugation and acidified to pH 4–5 with 10% HClO\(_4\) to hydrolyze the 5’-imidazolide groupings (Schwartz and Orgel, 1985). For HPLC analysis, 40 \(\muL\) of the reaction mixture was injected onto the anion exchange column, and trimer fractions were collected and stored at \(20^\circ\)C until further use. To obtain sufficient amounts of trimer fraction, several such injections were made.

### Isolation of linear isomers from trimer fraction

As the anion exchange column separates the mixture according to the number of negative charges on each component, collected trimer fractions contain linear isomers along with cyclic isomers and pyrophosphate derivatives. To isolate the linear isomers from this mixture, the trimer fraction was hydrolyzed with APH, an enzyme that cleaves exo-phosphate groups yet is inactive TABLE 1. THE PRODUCTS FORMED FROM RNase T2 HYDROLYSIS FOLLOWED BY APH HYDROLYSIS OF EACH TRIMER FRACTION OBTAINED FROM THE PARTIAL HYDROLYSIS OF POLY[A,U]

<table>
<thead>
<tr>
<th>Starting isomer</th>
<th>RNase T2 hydrolysis product</th>
<th>APH hydrolysis product</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A^3pA^3pA)</td>
<td>(A^3p + A^3p + A)</td>
<td>(A + A + A)</td>
</tr>
<tr>
<td>(A^3pA^3pU)</td>
<td>(A^3p + A^3p + U)</td>
<td>(A + A + U)</td>
</tr>
<tr>
<td>(A^3pU^3pU)</td>
<td>(A^3p + U^3p + U)</td>
<td>(A + U + U)</td>
</tr>
<tr>
<td>(U^3pA^3pA)</td>
<td>(U^3p + U^3p + A)</td>
<td>(U + U + A)</td>
</tr>
<tr>
<td>(U^3pA^3pU)</td>
<td>(U^3p + A^3p + U)</td>
<td>(U + A + U)</td>
</tr>
<tr>
<td>(U^3pU^3pU)</td>
<td>(U^3p + U^3p + U)</td>
<td>(U + U + U)</td>
</tr>
</tbody>
</table>

Linear trimers with the general formula of \(NpNpN\), where \(N = A\) or \(U\) for \(\text{ImpA-ImpU}\) reaction, were collected and further separated by reverse phase HPLC (Fig. 2).

### Fraction collection and enzymatic hydrolysis

For enzymatic hydrolysis studies, fractions were collected, either individually or two together—depending upon the difference between their close retention times, fractions 4 and 5; 6 and 7 and 8 were collected together. They produced different hydrolysis products.

FIG. 2. Reverse phase HPLC elution profile of linear trimer isomers formed in the montmorillonite-catalyzed reaction of ImpA with ImpU. Alltima C-18, 5\(\mu\)L, 4 \(\times\) 250 mm reverse phase column. Mobile phase: Buffer A: 0.02 \(M\) NaH\(_2\)PO\(_4\) in 0.2% w/v trifluoroacetic acid solution at pH 2.5; Buffer B: 0.2% trifluoroacetic acid, pH 2.5 in 30% acetonitrile v/v. Flow rate: 1.0 mL/min. Gradient: Initial time to 10 min: Buffer B 0 to 15%; 10–14 min: Buffer B 15%; 14–29 min: Buffer B 15 to 30%; 29–33 min: Buffer B 30 to 40% (HPLC conditions: Kanavarioti, 1997). Because of their close retention times, fractions 4 and 5; 6 and 7 and 8 were collected together. They produced different hydrolysis products.
Sequence analysis of trimer isomers—Their retention times—as indicated in Fig. 2. Collected fractions were first hydrolyzed with RNase T2 followed by APH. Hydrolysis products were identified by co-injections with authentic standards on reverse phase HPLC.

The concentration of each fraction to be hydrolyzed with RNase T2 or APH was calculated from its UV absorbance, which was measured when it was re-injected onto reverse phase or anion exchange column to check its purity. This value was then used to calculate the amount of enzyme to be added to the hydrolysis mixture using the instruction sheet supplied by the manufacturer. A schematic that illustrates the complete analysis procedure is given in Fig. 3.

Sequence identification of isomers was accomplished by combining the data obtained from their reverse phase HPLC elution orders and from the identification of the enzymatic hydrolysis products that were co-injected with authentic standards. Results are listed in Table 2.

Results

There are $2^4 = 32$ possible linear trimer isomers that can be formed from the binary mixtures of monomers: 2 kinds of monomer and 4 possibilities for bond formation, namely, $pN^2\ pN^2\ pN$, $pN^2\ pN^3\ pN$, $pN^3\ pN^2\ pN$, and $pN^3\ pN^3\ pN$. We identified the sequences of 16 isomers (Table 2) and found that 88% of these linear trimers were hetero-isomers, and only 12% were in the form of the 8 possible homo-trimers. In a random distribution, 25% of the trimers would be expected to be homo-isomers ($8/32$). Overall incorporation of monomer units into trimer isomers was found to be about 61 and 39% for A and U, respectively, based on the yields calculated from the area ratios after normalization to 100. As shown in Table 3, 61% of the monomers were linked via RNA-like $3',5'$-phosphodiester bonds.

Discussion

The ultimate goal of our work is to form RNA-like oligomers (i.e., hetero-oligomers with sequence specificity) using minerals as a catalyst. In contemporary RNA, with very few exceptions, all of the monomer units are linked via $3',5'$-phosphodiester bonds. Genetic information encoded in DNA is copied and passed from mother- to daughter-DNA strands by replication and from DNA to RNA strands by transcription. Both transcription and replication processes are catalyzed by enzymes. Although RNA-like oligomers formed by montmorillonite catalysis also contain $2',5'$-linkages, it has been shown that synthetic oligocytidylates formed by montmorillonite catalysis and that contain more than 50% $2',5'$-linkages, or exclusively $2',5'$-linkages, do serve as an efficient catalyst to form the complementary oligoguanylates without the aid of enzymes. Furthermore, the monomer units in the product oligoguanylate strands formed on exclusively $2',5'$-oligocytidylate chains contain $2',\ 5'$-, as well as $3',5'$-linkages (Ertem and Ferris, 1996).

One of the major concerns in studies of abiotic formation of bio-oligomers is the formation of the

---

**FIG. 3.** A schematic illustrating the complete analysis procedure. AE-HPLC, Anion Exchange High Performance Liquid Chromatography; RP-HPLC, Reverse Phase High Performance Liquid Chromatography.
phosphodiester bond (or peptide bond in proteins) in very dilute aqueous solutions, which is an equilibrium reaction that requires removal of water molecules from the medium to shift the equilibrium in favor of product formation. Our work has demonstrated that mineral catalysis may have overcome this obstacle by facilitating the formation of the phosphodiester bond between monomers in dilute aqueous solutions, thereby producing RNA-like hetero-oligomers.

The proportion of RNA-like 3', 5'-linkages observed for the dimer isomers (Table 1 in Ertem and Ferris, 2000), which was 51%, increased to 61% in trimer isomers, while the ratio of hetero-isomers increased from 73% to 88% for dimer and trimer isomers, respectively.

Sequence analysis of trimers, tetramers, and pentamers formed in the montmorillonite-catalyzed ImpA-ImpC reaction has been reported (Miyakawa and Ferris, 2003). In our reaction, the trimer and tetramer yields were found to be 14% and 6.7%, respectively, based on the area ratios of the peak of interest to the total area on the chromatogram. Linear isomers constituted 67% of the trimer fraction. Our attempts to determine the sequences of oligomers longer than trimers (i.e., tetramers and pentamers) have failed to produce reliable results because they form in low yields and produce a large number of isomers. Further, we lacked authentic standards and a suitable column to resolve the isomers.

**CONCLUSIONS**

Na-montmorillonite, one of the most abundant phyllosilicate-group minerals on Earth, which have also been identified on Mars, serves as a catalyst for the formation of phosphodiester bonds in an aqueous electrolyte solution and produces RNA-like oligomers. Our analysis of the trimer fractions formed in the reaction of binary mixtures of activated monomers demonstrated that montmorillonite catalysis facilitated the formation of hetero-oligomers with sequence selectivity similar to that of short RNA segments: 88% of the linear trimers were RNA-like hetero-isomers that contained 61% A-monomer and 39% U-monomer, and 61% of the phosphodiester bonds were RNA-like 3', 5'-linkages.

With the currently available analytical tools and methodology, we have been able to identify the se-
sequences of only 16 isomers, which constitute 83% of the linear trimers. The number and the sequences of the isomers contained in the remaining 17% of the linear trimer fraction are, therefore, currently unknown. Contrary to the general trend that favors the formation of a limited number of isomers (Ertem and Ferris, 2000; Miyakawa and Ferris, 2003), the ideal scenario would be to form all of the 32 isomers, which, in turn, would provide a richer pool for the selection of “useful, functional” isomers in the course of the evolution.

Our findings of 88% hetero-isomer formation are, therefore, highly significant for origin of life studies. These model studies lend support to Bernal’s hypothesis that mineral catalysis may have played an important role for the abiotic formation of bio-molecules in the early stages of chemical evolution.

ACKNOWLEDGMENTS

This and our previous research on the mineral-catalyzed formation of RNA-like oligomers were inspired by the pioneering works of Professor Leslie E. Orgel of the Salk Institute for Biological Studies on the oligomerization of activated monomers and non-enzymatic template-directed synthesis of oligonucleotides. Most of the experimental procedures carried out in this work have been published in his numerous publications. We are grateful for the most valuable knowledge we have gained from his work over the years.

We are also grateful for funding support from the NASA Astrobiology Institute via the Carnegie Institution of Washington and the NASA Goddard Center for Astrobiology, and for a National Research Council Senior Fellowship for GE.

ABBREVIATIONS

APH, alkaline phosphatase enzyme; HEPES, N-(2-Hydroxyethyl)piperazine-N’-(2-ethanesulfonic acid). HPLC, High Performance Liquid Chromatography; ImpA, 5’-phosphorimidazolide of adenosine; ImpC, 5’-phosphorimidazolide of cy-
REFERENCES


Shapiro, R. (2006) Small molecule interactions were central to the origin of life. Q. Rev. Biol. 81, 105–125.