ACS OMEGA

Open Access

This article is licensed under CC-BY-NC-ND 4.0

http://pubs.acs.org/journal/acsodf

Article

A Prebiotic Genetic Nucleotide as an Early Darwinian Ancestor for Pre-RNA Evolution

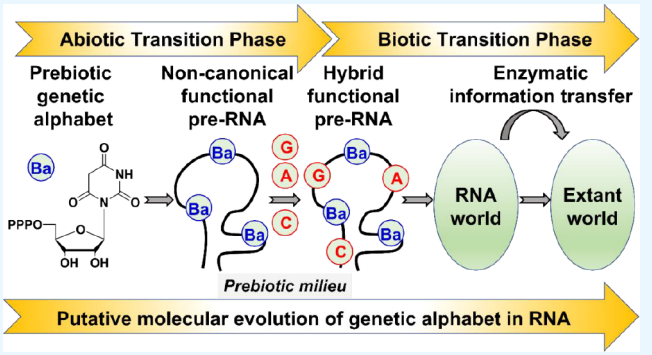
Anupam A. Sawant,* Sneha Tripathi, Sanjeev Galande, and Sudha Rajamani*

Cite This: *ACS Omega* 2024, 9, 18072−18082

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Prebiotic genetic nucleotides (PGNs) often outcompete canonical alphabets in the formation of nucleotides and subsequent RNA oligomerization under early Earth conditions. This indicates that the early genetic code might have been dominated by pre-RNA that contained PGNs for information transfer and catalysis. Despite this, deciphering pre-RNAs' capacity to acquire function and delineating their evolutionary transition to a canonical RNA World has remained under-researched in the origins of life (OoL) field. We report the synthesis of a prebiotically relevant nucleotide (BaTP) containing the noncanonical nucleobase barbituric acid. We demonstrate the first instance of its enzymatic incorporation into an RNA, using a T7 RNA polymerase. BaTP's incorporation into baby spinach aptamer allowed it to retain its overall secondary structure and function. Finally, we also demonstrate faithful transfer of information from the pre-RNA-containing BaTP to DNA, using a high-fidelity RNA-dependent DNA polymerase, alluding to how selection pressures and complexities could have ensued during the molecular evolution of the early genetic code.



INTRODUCTION

Life is thought to have emerged on the early Earth around 3.5b years ago.[1−3] Two approaches that have dictated the origins of life studies include a top-down "biology-centric" approach and a bottom-up "chemistry-inspired" approach.[4] The former is based on phylogenetic analysis to deduce the simplest cellular life form called LUCA (the Last Universal Common Ancestor), which is thought to be a compartmentalized organism with a complex metabolic and catalytic network, along with a sophisticated genetic system.[5,6] The bottom-up approach relies on studying the emergence of life from simpler interacting organic chemicals.[7−9] Understanding the chemical emergence of life poses the following two prominent questions: (1) which biomolecule came about first?[10] (2) what and how many steps were potentially involved in this process that led to the emergence of the first cellular life? Various models have been proposed, including protein-first,[11,12] metabolism-first,[13,14] and membrane-first models,[15,16] which have been fundamental to understanding life's emergence.

Among these multipronged approaches, there is the much-studied nucleic acid-first model that builds on RNA's dual ability to transfer information and perform catalysis.[17−20] Pertinently, the importance of informational molecules in driving Darwinian evolution further adds weight to this model. The fundamental step in RNA-based evolution is the formation of nucleotide monomers and their subsequent oligomerization under early Earth conditions. This is then thought to have been followed by templated replication to enrich for informational oligomers, some of which then adopt secondary structures and acquire functions, eventually establishing a putative RNA World.[21] Noteworthy is that when the constituents of a nucleoside monomer are allowed to react under prebiotic conditions, it almost always results in the "nucleosidation problem," i.e., the difficulty faced during the condensation of a nucleobase onto the sugar.[22−25] Therefore, extant nucleobases might not seem that "special" if one aims to study RNA evolution based mainly on intrinsic reactivity.

A natural respite to the aforesaid issue is provided by the heterogeneity inherent to the prebiotic soup, as it is a diverse mixture of various heterocycles that, in principle, have the potential to result in nucleobases.[26−28,37] In this context, earlier studies have explored the possibility of the formation of nucleo(si)tide monomers or PGNs under pertinent early Earth conditions. Strikingly, the noncanonical nucleobases evaluated seem to fare better in comparison to extant nucleobases. The above research findings essentially demonstrated that the nucleosidation problem can be circumvented by invoking

Received: December 13, 2023
Revised: March 22, 2024
Accepted: March 29, 2024
Published: April 12, 2024

© 2024 The Authors. Published by American Chemical Society

prebiotically relevant nucleobases or PGNs.[29−35] Noteworthy is that many of these PGNs are less prone to depurination, which results during RNA oligomerization under early Earth geochemical conditions.[36] Most importantly, post glycosylation reactions, the PGNs often form analogs that are isomorphic to extant nucleo(si)tides in terms of their H-bonding faces.[37] Despite the aforementioned features and the propensity of these prebiotic nucleoside/tide surrogates to result in intact informational oligomers (pre-RNAs), a systematic understanding of the ability of PGN containing pre-RNAs, to adopt structures essential to emerging as functional RNAs, such as aptamers and ribozymes, is lacking. Aptamers are functional RNA motifs that have the propensity to fold into distinct three-dimensional structures that result in binding to selected ligands with very high affinities.[38−40] Parallelly, the biology-centric approach has seen various efforts being put into creating artificial or semisynthetic informational polymers to expand the genetic code of extant life.[41,42] The artificial genetic alphabets used in these studies typically comprise modified nucleobases that have been used to study base-pairing properties that they exhibit, as well as to characterize the fidelity of information transfer using both *in vitro* and *in vivo* approaches.[43−45] These studies also raise an important question concerning the genetic code of life and its possible molecular evolution during the genetic transition from ancient to extant life. Further, the use of several modified nucleobases in extant biology adds compelling weight to the pre-RNA World conjecture.[46]

In this aforementioned context, PGNs containing noncanonical nucleobases that are isomorphic to contemporary pyrimidine nucleobases have shown promise as candidates for studying RNA-based evolution during life's early stages.[47,48] Characterizing the formation and emergent functions of these early RNA molecules would be central to understanding the evolution of the genetic system(s) *per se*. Against this backdrop, we report the synthesis of one such putative PGN triphosphate that contains a barbituric acid nucleobase (namely, Barbitudine triphosphate or BaTP). Barbituric acid (Ba) is a prebiotically pertinent noncanonical heterocyle whose presence has been demonstrated in meteorites and under prebiotically relevant reaction conditions.[37] Further, we demonstrate the incorporation of this BaTP into RNA, using a bacterial T7 RNA polymerase. To our knowledge, this is the first report of this particular moiety's incorporation via an enzymatic route. Further, we show that the incorporation of BaTPs into a functional baby Spinach aptamer (bSP) RNA does not hamper the folding of the resultant aptameric RNA. This emphasizes the suitability of BaTP-containing aptamer RNA for acquiring ligand binding functions, which we also demonstrate by studying the binding of this aptamer to its ligand DFHBI. Lastly, we also explored the propensity for information transfer from this PGN-containing hybrid RNA into DNA, by subjecting it to reverse transcription and PCR amplification. This is not only important for discerning the faithful transfer of genetic information from the PGN-containing pre-RNA to DNA but also has implications for delineating the transition of the genetic code from the pre-RNA world(s) to the RNA World. Subsequent molecular evolution driven by other selection pressures and constraining environmental parameters is considered to have eventually resulted in the molecular complexity that we presently see in the extant World.

MATERIALS AND METHODS

Reagents. Trimethylsilyl trifluoromethanesulfonate, *O*-bis(trimethylsilyl)acetamide, Barbituric acid, 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-beta-D-ribofuranose, Sodium methoxide, and DOWEX 50 × 8 H^+ resin were all procured from Sigma-Aldrich. DNA oligonucleotides were purchased from Eurofins, Inc. and were used without further purification. T7 RNA polymerase, ribonuclease inhibitor (RiboLock), NTPs, RNase A, and RNase T1 were obtained from Fermentas Life Science. Calf intestinal alkaline phosphatase (CIP) and snake venom phosphodiesterase I were procured from Invitrogen (ThermoFisher Scientific) and Sigma-Aldrich, respectively. Chemicals for preparing buffer solutions were purchased from Sigma-Aldrich (BioUltra grade). Autoclaved water was used in all the biochemical reactions, and for HPLC analyses, HPLC grade solvents were used.

Instrumentation. NMR spectra of small molecules were recorded on a 400 MHz Jeol ECS-400 and Bruker AVANCE III HD ASCEND 400 MHz spectrometers and processed using Mnova software from Mestrelab Research. Mass analysis was performed using an ESI-MS Waters Synapt G2-Si Mass Spectrometry instrument. HPLC analysis was done using Prominence UFLC Next Generation High-Throughput-Shimadzu HPLC using reverse phase shim-pack GIST C18 column (250 mm × 4.6 mm, 5 μm). The absorption spectra were recorded on a Shimadzu UV-2600 Shimadzu spectrophotometer. Steady-state and time-resolved fluorescence spectra of the control and modified spinach aptamers were recorded on a Fluoromax-4 spectrophotometer (Horiba Scientific). CD analysis was performed on a JASCO J-815 CD spectrometer. UV−Tm measurements were performed on a Cary 300 Bio UV−vis spectrophotometer.

Synthesis of 1-(2′,3′,5′-Tri-*O*-benzoyl) uridine 2 and Barbitudine. To the suspension of barbituric acid heterocyclic base (0.300 g, 2.34 mmol,1 equiv) in 15 mL of 1,2-dichloroethane under a nitrogen atmosphere, *N*,*O*-bis-(trimethylsilyl) acetamide (1.89 mL, 7.73 mmol, 3.3 equiv) was added. This reaction mixture was stirred until a clear solution was obtained. After this step, the powder of activated molecular sieves was added to the solution, followed by the addition of 1-*O*-acetyl-2,3,5-tri-*O*-benzoylribofuranose (1.05 g, 2.09 mmol, 0.9 equiv). This mixture was then allowed to stir for 15 min. After this, TMSOTf (200 uL, 1.16 mmol, 0.5 equiv) was added, and the mixture was heated under nitrogen and refluxed at 90 °C for 4 h. When the starting material was depleted significantly, the mixture was cooled to room temperature and the clear yellow solution was diluted with 50 mL DCM. The organic layer was then extracted with 50 mL ice-cold saturated $NaHCO_3$ solution and washed three times with 20 mL water. Organic layers were then dried over Na_2SO_4 and evaporated to dryness. The traces of impurities were removed by column chromatography to obtain compound **2** as a yellow-white solid (0.420 g, 32%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.02 (s, 1H), 8.04 (d, J = 7.5 Hz, 2H), 7.96 (d, J = 7.5 Hz, 2H), 7.85 (d, J = 7.6 Hz, 2H), 7.56−7.46 (m, 3H), 7.33 (dd, J = 31.9, 7.1 Hz, 6H), 6.42 (s, 1H), 6.05 (d, J = 10.6 Hz, 2H), 4.77 (d, J = 11.6 Hz, 1H), 4.67 (t, J = 8.0 Hz, 1H), 4.58 (dd, J = 11.8, 5.6 Hz, 1H), 3.69 (s, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 166.47 (s), 165.79 (s), 165.26 (d, J = 38.9 Hz), 164.94 (s), 164.24 (s), 149.66 (s), 133.82−133.03 (m), 133.16 (s), 130.07−129.80 (m), 129.68 (s), 128.89 (d, J = 8.8 Hz), 128.50 (d, J = 8.5 Hz), 86.97 (s), 79.26 (s), 74.02 (s),

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

71.11 (s), 63.77 (s), 39.78 (s). HRMS: (m/z). Expected mass for $C_{30}H_{25}N_2O_{10}$ $[M + H]^+$ = 573.1509, observed mass = 573.1506.

Finally, to obtain barbitudine, benzoylated nucleoside derivative **2** (0.140 g, 0.244 mmol, 1 equiv) was dissolved in 2 mL of absolute methanol, and to this solution, $NaOCH_3$ (0.012 g, 0.220 mmol, 0.9 equiv) was added. The reaction mixture was stirred at RT for 1 h, and the completion of the reaction was determined by TLC (dichloromethane:methanol = 3:1). Neutralization of the solution was achieved by adding a DOWEX-50 WX-8 ion-exchange resin (proton form) that was previously washed with methanol. After filtering the resin, the methanol was evaporated, and 20 mL of water was added. The benzoic acid methyl ester was extracted three times with 30 mL of diethyl ether each time, and the water fraction was lyophilized to give the deprotected nucleoside barbitudine as a pale-yellow solid (0.060 g, 94%). ^{1}H NMR (400 MHz, DMSO) δ 11.39 (s, 1H), 5.89 (s, 1H), 4.35 (s, 1H), 4.06 (t, J = 6.5 Hz, 1H), 3.66 (t, J = 7.7 Hz, 3H), 3.59 (dd, J = 11.6, 3.1 Hz, 2H), 3.39 (dd, J = 11.7, 6.5 Hz, 2H). ^{13}C NMR (101 MHz, DMSO) δ: 166.46 (d, J = 69.3 Hz), 151.0 (s), 87.86 (s), 84.17 (s), 71.43 (s), 69.94 (s), 62.27 (s), 40.01 (d, J = 99.2 Hz). HRMS: (m/z) Expected mass for $C_9H_{12}N_2O_7Na$ $[M + Na]^+$ = 283.0542, observed mass = 283.0547. This solid was used further without any purification step.

Synthesis of 6-Oxo Uridine-5′-triphosphate (BaTP). An ice-cold solution of the prebiotic ribonucleoside barbitudine (50 mg, 0.27 mmol, 1.0 equiv) in trimethyl phosphate (1 mL) was slowly added to freshly distilled $POCl_3$ (37 μL, 0.40 mmol, 1.5 equiv). The reaction mixture was stirred for 15 h at ∼4 °C. A solution of bis-tributylammonium pyrophosphate (0.5 M in DMF, 2.75 mL, 5.1 equiv) and tributylamine (0.65 mL, 2.75 mmol, 10 equiv) was rapidly added under ice-cold conditions. The reaction was quenched after 30 min with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.6, 15 mL) and was extracted with ethyl acetate (2 × 10 mL). The aqueous layer was evaporated and the residue was purified, first on a DEAE Sephadex-A25 anion exchange column (10 mM−1 M TEAB buffer, pH 7.6), followed by reversed-phase flash column chromatography (C18 RediSep Rf, 0−40% acetonitrile in 100 mM triethylammonium acetate buffer, pH 7.5, 50 min). Appropriate fractions were lyophilized to separate the desired triphosphate product BaTP as a (tetra)triethylammonium salt (42 mg, 16%). ^{1}H NMR (400 MHz, D2O) δ 6.15 (d, J = 2.8 Hz, 1H), 4.68 (dd, J = 6.3, 3.1 Hz, 2H), 4.42 (t, J = 6.7 Hz, 1H), 4.22 (dd, J = 8.5, 5.7 Hz, 1H), 4.08−4.01 (m, 2H). ^{13}C NMR (101 MHz, D2O) δ 179.73 (s), 166.67 (s), 87.38 (s), 81.15 (s), 71.72 (s), 69.46 (s), 65.84 (s), 58.66 (s). ^{31}P NMR (162 MHz, D2O) δ −10.58 (dd, J = 49.6, 19.8 Hz), −22.83 (s). HRMS: (m/z) Expected mass for $C_9H_{14}N_2O_{16}P_3$ $[M−H]^−$ = 498.9556, observed mass = 498.9553.

In Vitro Transcription of Baby Spinach DNA Template in the Presence of BaTP and UTP. The transcription reactions were carried out in 40 mM Tris-HCl buffer (pH 7.9) containing 200 ng of the DNA template (see DNA template design below), 1 mM GTP, CTP, ATP, UTP/BaTP, 10 mM $MgCl_2$, 10 mM NaCl, 10 mM dithiothreitol (DTT), 2 mM spermidine, 1 U/μL RNase inhibitor (Riboblock), and 40 units of T7 RNA polymerase, in a total volume of 25 μL. Samples were incubated overnight at 37 °C. The RNA products were purified by phenol-chloroform extraction, and the RNA pellet formed from each reaction was washed with 70% aqueous ethanol, dried, and dissolved in 20 μL of nuclease-free water. Control (UTP-containing) and modified (BaTP-containing) transcripts obtained after transcription reactions were resolved and detected by 2% agarose gel containing EtBr. Further, the presence of Barbitudine in the RNA transcript was confirmed by enzymatic digestion followed by HPLC analysis of the ribonucleoside products obtained from the digestion reaction (For experimental details, please see enzymatic digestion protocol described below in the relevant section).

Details of DNA Template Design. The regions underlined depict the binding sites for the reverse primer and the forward primer.

5′TAA TAC GAC TCA CTA TAG GGT GAA GGA CGG GTC CGT TCG CGT TGA GTA GAG TGT GAG CTC C 3′

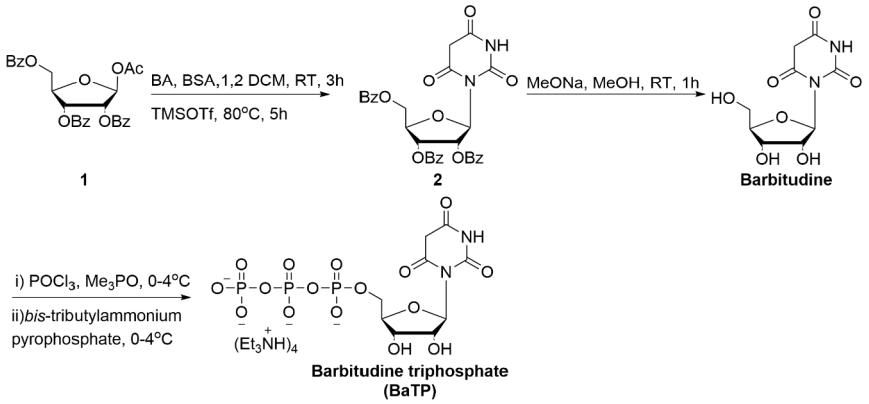
Enzymatic Digestion of BabSP. Approximately four nanomoles of only the modified oligoribonucleotide transcripts that were obtained after the transcription reaction, which was carried out either in the presence of only BaTP or with equimolar concentration of BaTP:UTP, was taken for digestion experiment(s). These were then treated (separately) with snake venom phosphodiesterase I (0.01 U), calf intestinal alkaline phosphatase (10 μL, 1 U/μL), and Rnase A (0.25 μg), in a total volume of 100 μL in 50 mM Tris-HCl buffer (pH 8.5, 40 mM $MgCl_2$, 0.1 mM EDTA) for 12 h at 37 °C. After this period, Rnase T1 (0.2 U/μL) was added, and the samples were incubated for another 4 h at 37 °C. The ribonucleoside mixtures obtained from the digests were analyzed by reversed-phase HPLC using the Shimadzu Shim-pack GIST C18 column (250 × 4.6 mm, 5 μm) at 260 nm. Mobile phase A: 50 mM triethylammonium acetate buffer (pH 7.5); mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min.

CD Analysis of Control (UbSP) and Modified (BabSP) Aptamer RNAs. For CD analysis, solutions of 5 μM control baby spinach aptamer RNA (UbSP, which had **U**-A-G-C) and modified aptamer RNA (BAbSP, which had Ba-A-G-C) were prepared in sensor buffer containing 10 mM Tris at pH 7.5, 50 mM KCl, and 5 mM $MgCl_2$, in nuclease-free water. The samples were vortexed and spun down, annealed for 4 min at 90 °C in a heating block, and subsequently allowed to come to RT. CD analyses were carried out soon thereafter, and the CD spectra were recorded on a Jasco J-815 spectrometer using a standard quartz cuvette with a 1 cm path length. Scans from 200 to 320 nm were acquired with a scanning speed of 100 nm/min, a step of 1 nm, and a bandwidth of 1 nm. A positive band at 264 nm confirmed the formation of a G-quadplex binding pocket, as previously reported by DasGupta et al. for G-quadruplex-containing RNAs, including the spinach aptamer derivatives.[61] Similar CD spectra obtained for the control and modified baby spinach RNAs confirmed the formation of the G-quadruplex binding pocket in modified baby spinach RNA (BAbSP).

UV-Thermal Melting (Tm) of Control (UbSP) and Modified (BabSP) Aptamer RNAs. For UV−Tm analysis, solutions of 5 μM control baby spinach aptamer RNA (UbSP, which had U-A-G-C) and modified aptamer RNA (BAbSP, which had **Ba**-A-G-C) were prepared in sensor buffer containing 10 mM Tris pH 7.5, 50 mM KCl, and 5 mM $MgCl_2$, in nuclease-free water. The samples were vortexed and spun down, annealed for 4 min at 90 °C in a heating block, and subsequently allowed to come to RT. Samples were incubated

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

Scheme 1. Synthesis of BaTP



Scheme depicting various steps in the chemical synthesis of barbitudine and its corresponding triphosphate, i.e., barbitudine triphosphate (BaTP)

at RT for 1 h and used for thermal melting analysis. Measurements were performed on a Cary 300 Bio UV−vis spectrophotometer. Temperature was increased from 20 to 90 °C at 1 °C min^{-1} interval, and the absorbance was measured every 0.5 °C at 260 nm.

Fluorescence Analysis of Control (UbSP) and Modified (BabSP) Aptamer RNAs. RNA samples (5 μM) were prepared for fluorescence analysis using a protocol similar to that described for CD analysis. Additionally, a DFHBI (10 μM) ligand was added to the UbSP and BAbSP aptamer RNA samples, respectively. In addition to these, a DFHBI control (minus the aptamer RNA) was also prepared in the exact same manner to account for any intrinsic DFHBI fluorescence. Before fluorescence analysis, RNA samples were incubated in the dark for 2 h. Following incubation, fluorescence analysis was recorded on a Fluoromax-4 spectrophotometer (Horiba Scientific) using excitation at 460 nm and emission at 500−505 nm. The slit width used for excitation and emission was 8/8, respectively. All of the respective controls were also analyzed using similar instrumental parameters.

Reverse Transcription and PCR Amplification of Barbitudine-Modified mCherry mRNA. One μg of mCherry mRNA (1 μg) was taken for cDNA synthesis using Superscript III first strand synthesis kit (Invitrogen catalog no. 18080051). The cDNA synthesis was done as per the manufacturer's recommended protocol. Briefly, the mCherry RNA, random hexamers, and dNTPs were incubated at 65 °C for 5 min and immediately put on ice (4 °C) for another 5 min. The 5× first strand synthesis buffer was added to the above mix along with 0.1 M DTT, RNase Inhibitor, and the Superscript III enzyme. The above mixture was subjected to the following incubations: 25 °C for 5 min, 50 °C for 30 min, followed by 70 °C for 15 min. cDNA validation was performed by using HiFi PFU enzyme (DX/DT catalog no. R1220) using standard PCR protocol. Briefly, the PFU master mix (2×) was added to the forward−reverse primers (see details below) and the cDNA template, and the PCR was set up as per the recommended protocol.

Details of mCherry DNA Template Design. The regions underlined depict the binding sites for the reverse primer and the forward primer.

5′ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGTAG3′

Forward Primer. 5′TAATACGACTCACTATAGGGGATGGTGAGCAAGGGCGAGGAG3′

Reverse Primer. 5′CTTGTACAGCTCGTCCATGCC3′

RESULTS AND DISCUSSION

Synthesis of Barbitudine, a PGN Containing Barbituric Acid. Barbituric acid was recently explored as a prebiotically pertinent nucleobase that readily forms nucleoside and nucleotide analogs under prebiotically relevant conditions.[36,48] Further, the nucleoside monophosphate of barbituric acid was observed to oligomerize effectively under strongly acidic conditions at high temperatures, whereas the oligomerization of canonical nucleobases under these same conditions has been shown to result in abasic oligomers.[49] Abiotically formed barbituric acid nucleoside/tide is isomorphic to uridine and pseudo uridine (ubiquitously found in structural RNAs),[50] making barbitudine an interesting PGN candidate to study early RNA evolution. Additionally, Krishnamurthy[51] and Powner[52] groups recently also described a putative route for the prebiotic formation of nucleoside triphosphates, molecules that are thought to have paved the way for the transition from abiotic to biotic. Given all of these aforementioned promising studies, we felt encouraged to synthesize barbitudine triphosphate (BaTP) and systematically explore its use as a putative pre-RNA World PGN.

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

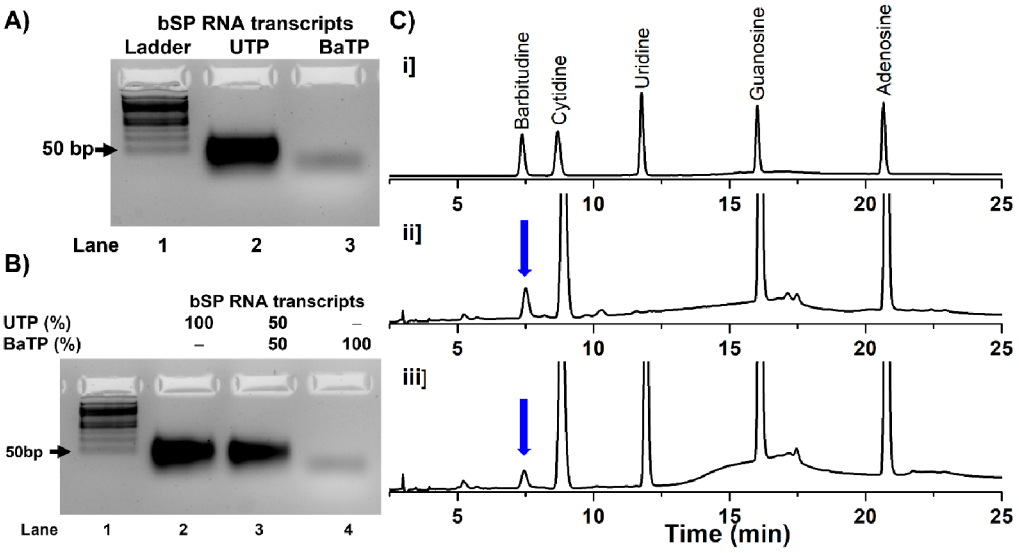


Figure 1. Enzymatic incorporation and digestion of BaTP-modified RNA. (A) 3% agarose gel showing the baby spinach (bSP) RNA product that resulted from in vitro transcription (IVT) reactions performed in the presence of UTP and the prebiotic genetic nucleotide, BaTP. Lane 1, 50 bp ladder. Lane 2 shows the RNA transcripts from the IVT reaction, which used the four canonical nucleotides, including UTP. Lane 3 shows the RNA transcripts from the IVT reaction where the UTP was totally replaced with BaTP. Comparison of lanes 2 and 3 indicates the formation of full-length RNA transcripts, which should be in the range of 40−50 bp. (B) The competitive incorporation of natural UTP versus the PGN BaTP by the T7 RNA polymerase enzyme was assessed by performing a reaction in the presence of equimolar concentrations of UTP and BaTP (lane 3). (C) HPLC chromatograms of ribonucleoside products obtained from the enzymatic digestion of RNA transcripts at 260 nm. (i) Standard mix of natural ribonucleosides and Barbitudine. (ii) RNA digests that were obtained from IVT reactions carried out in the presence of BaTP plus other 3 canonical NTs. (iii) RNA digest obtained from IVT reaction carried out in the presence of an equimolar concentration of UTP and BaTP plus other 3 canonical NTs (sample extracted from band in lane 3 of Figure 1B). HPLC details: mobile phase A: 100 mM TEAA (pH 7.5); mobile phase B: acetonitrile. Flow rate: 1 mL/min. Gradient: 0−10% B in 20 min and 10−100% B in 10 min.

We synthesized barbitudine using a combination of previously reported synthetic protocols[53,54] with few modifications added as required.[55] The nucleobase BA was made to react with commercially available acetate donor **1** in the presence of *N*,*O*-bis(trimethylsilyl)acetamide (BSA), using TMSOTf as a Lewis acid catalyst (Scheme 1). Barbituric acid has two nucleophilic centers in the form of imine nitrogens (N1 and N3) and carbanion carbon C-5, which is flanked by two carbonyl groups. This raised the possibility of the formation of two nucleosides; the C-5 linked nucleoside and the N1/N3-nucleoside as has been reported when glycosylation of BA on to ribose was performed using geologically relevant wet−dry cycles.[36] Our reaction conditions used for the chemical synthesis mainly yielded N1/N3 nucleosides as the major product, though we also saw a minor amount of C-5 nucleoside formation. The resultant protected ribonucleoside analog **2** was purified using normal-phase chromatography, which was then analyzed by HRMS and ^{1}H, ^{13}C, and ^{13}C DEPT NMR analyses to confirm the regiochemistry (N-glycoside vs C-glycoside) of the nucleoside (Figures S1−S4). ^{1}H NMR showed two proton integrations for the barbituric acid nucleobase at the C-5 position. Additionally, ^{13}C DEPT showed a negative peak for the same C-5 carbon of barbituric acid, indicating two proton carbons. Both ^{1}H and ^{13}C NMR confirmed the formation of the N-glycoside.

Further, this protected ribonucleoside **2** was subjected to deprotection in sodium methoxide and methanol to obtain the free ribonucleoside barbitudine. This was further extracted in water and washed with diethyl ether to remove the methyl ester of benzoic acid. The aqueous layer was lyophilized to yield a pure barbitudine nucleoside (Figures S5−S8), which was used for further experiments. Since the substrate required for the enzymatic incorporation is a triphosphate of barbitudine, it was successfully converted into barbitudine triphosphate (BaTP). This was done by a one-pot reaction using $POCl_3$, followed by treatment with bis-pyrophosphates in trimethyl phosphate as a solvent (Scheme 1). The resultant BaTP was systematically purified, first using anion exchange chromatography, followed by reverse phase chromatography. The final compound was then subjected to HRMS and NMR analysis to confirm the product's integrity (Figures S9−S12).

Enzymatic Incorporation of BaTP into Baby Spinach (bSP) RNA Aptamer. There are two routes by which BaTP could be incorporated into an RNA to study putative emergent properties of such pre-RNA containing polymers. One is by using a chemical route that works for short RNAs, while the other involves using the enzymatic route. Since we also aimed to explore the suitability of barbitudine to demonstrate putative abiotic to biotic transition, we proceeded with the enzymatic incorporation route. Pertinently, this also provides a "natural" selection pressure that allows for gauging BaTP's incorporation into functional length RNA stretches. This, in turn, could also shed light on the feasibility of molecular evolution that is thought to have potentially occurred during early genetic transitions. Another important aspect of *in vitro* RNA synthesis using T7 RNA polymerase and DNA templates is the formation of multiple shorter RNA transcripts along with the intended full-length RNA transcript (and transcripts with up to +3 additions).[56] This scenario provides a realistic backdrop for characterizing emergent properties of RNA(s) from a pool of varying lengths of RNA populations, which is also reflective of what could have transpired in a heterogeneous prebiotic soup.

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

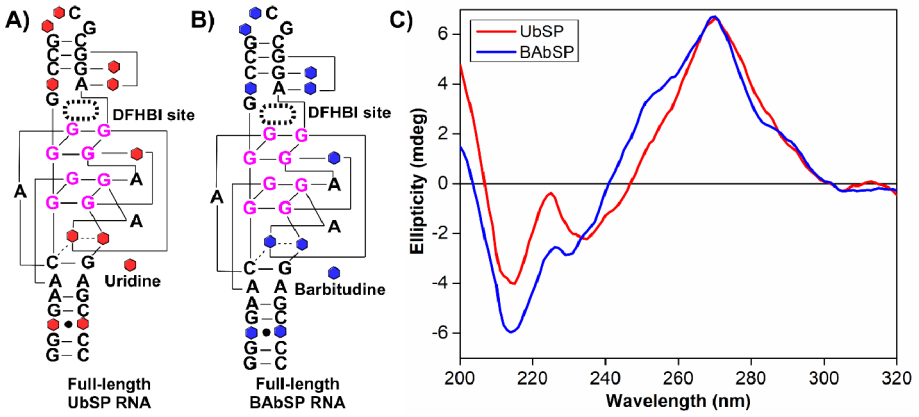


Figure 2. Secondary structure of UbSP and BAbSP aptamer RNAs. CD spectra of 5 μM unmodified UbSP (panels A and C, red trace) and modified BAbSP (panels B and C, blue trace) RNA, respectively. Both RNA aptamer samples were prepared in a sensor buffer containing 10 mM Tris-HCl, 100 mM KCl, 10 mM $MgCl_2$, and at pH 7.4.

In this backdrop, we explored *in vitro* transcription reactions for incorporating BaTP into a functional RNA motif.[57] We chose a baby spinach DNA template, which typically results in a 43-nt long functional RNA aptamer (bSP aptamer) after undergoing a successful transcription reaction in the presence of the canonical UTP[58] (see DNA template design in the Materials and Methods section). In a parallel reaction, BaTP incorporation was assessed by performing the T7 RNA polymerase reactions in the presence of this PGN (along with A, U, and C). The baby spinach DNA template was designed to contain ten dA residues in the coding region. If the BaTP-based transcription reaction proceeded as expected, these dAs would guide the incorporation of the monophosphate of BaTP into the RNA transcripts.

In this backdrop, transcription reactions were performed in the presence of GTP, CTP, ATP, and UTP in the control RNA reaction or by replacing UTP with BaTP in the modified reaction. The resultant products were resolved using 3% agarose gels. Our studies showed that the modified transcription reaction that involved the use of BaTP, ATP, CTP, GTP, and the template DNA resulted in lesser yield. The gel analysis indicated the formation of a putative full-length RNA transcript along with some altered-length RNA transcripts when compared to the control reaction where the canonical UTP was used along with ATP, CTP, and GTP (Figure 1A, lane 3 vs lane 2). To further discern the outcome of the transcription reaction, we performed HPLC analysis of the RNA transcripts (Figure S13). The HPLC trace showed a similar population distribution in the case of both the unmodified (UbSP transcripts) and the modified RNA transcripts (BAbSP transcripts). This result was further corroborated by 20% denaturing PAGE analysis that also showed truncated transcript formation in both the UbSP and BAbSP RNA that resulted from the *in vitro* transcription reactions (Figure S14). Further, we also analyzed the control UbSP and modified BAbSP RNA transcripts by subjecting them to HRMS analysis to characterize the control and the modified RNA transcripts. Nonetheless, even after multiple attempts and trying to vary the mass analysis conditions, we were unable to obtain HRMS data for both the control and modified RNA transcripts. In addition to instrument-related sensitivity issues that made working with 43-mer length of RNA molecules very challenging, the issue got compounded due to the intrinsic complexity and the low yield of the resultant product mixtures (Figures S13 and S14).[59]

Therefore, to confirm the incorporation of barbitudine triphosphate into the bSP RNA aptamer, we performed enzymatic digestion of the RNA that was obtained from the transcription reactions containing BaTP-CTP-GTP-ATP (Figure 1A, lane 3), and the resulting RNA digest was subjected to HPLC analysis. The HPLC chromatogram from the digestion of BAbSP RNA did indeed show a peak for barbitudine. This nucleoside eluted at the same retention time as that of the corresponding peak that was observed in the chromatogram obtained for a mix of nucleoside standards (Figure 1C (i) and (ii) panels, respectively). We further assessed the competitive incorporation of BaTP versus the canonical UTP by performing an in vitro transcription reaction in the presence of 1:1 concentrations of UTP and BaTP, along with the other three canonical NTPs (Figure 1B). Further, the RNA digest of the transcript obtained from this competitive incorporation of BaTP and UTP (1:1), also showed a peak at a retention time similar to that of the barbitudine used in the standard mix (Figure 1C iii vs i panel).

The HPLC analysis of ribonucleoside products arising from enzymatic digestion of barbitudine-containing RNA transcripts from the aforementioned reactions clearly revealed the incorporation of the ribonucleoside into the RNA transcripts. We further confirmed the integrity of the natural as well as barbitudine ribonucleoside by subjecting the corresponding HPLC fraction to mass characterization (Table S1 and Figures S15−S18). These results further unambiguously ascertained the presence of the PGN barbitudine in the bSP RNA transcripts. Pertinently, these enzymatic digestion results highlight the qualitative ability of BaTP to compete with canonical UTP during transcription reactions.

Secondary Structure Analysis of Unmodified (UbSP) and Modified (BabSP) RNA Aptamers using CD Spectroscopy Analysis. As mentioned earlier, the transcription reaction results in a scenario that reflects a heterogeneous prebiotic soup in which RNAs that have the ability to form secondary structures could have been selected for in certain contexts, e.g., catalysis, ligand binding by aptamer, etc. The presence of such structures can be delineated using spectroscopic methods such as circular dichroism (CD), a very useful technique for characterizing RNA aptamers that contain G-quadruplex forming sequences. This is pertinent as

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

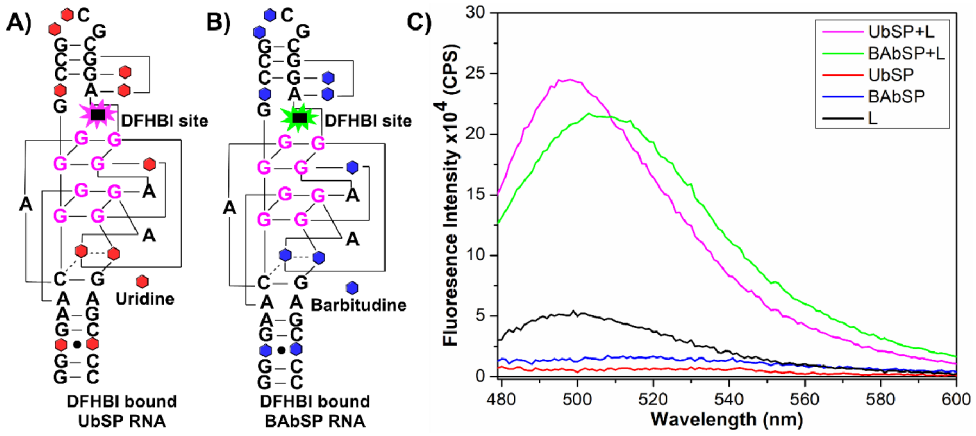


Figure 3. Fluorescence analysis of (A) UbSP and (B) BAbSP aptamer RNAs. Five μM UbSP and BAbSP aptamers were prepared in the sensor buffer containing 10 mM Tris-HCl, 100 mM KCl, and 10 mM $MgCl_2$ at pH 7.4. To this, 10 μM of the DFHBI ligand (L) was added and the mixture was further incubated for 2 h in the dark. (C) An evident increase in fluorescence intensity was observed for UbSP (pink trace) and BAbSP (green trace) RNAs in the presence of DFHBI. The control samples of only UbSP (minus DFHBI; red trace), only BAbSP (minus DFHBI; blue trace), and DFHBI-alone control (black trace) did not show any significant fluorescence (excitation at 460 nm and emission at 500−505 nm).

the G-quadruplex plays a crucial role in the secondary structure formation of UbSP and BAbSP RNA aptamers, which in turn impinges on its functional capacity to bind its ligand. Therefore, for CD analysis, 5 μM of unmodified (UbSP) and 5 μM of modified (BAbSP) RNAs were prepared in a sensor buffer containing 10 mM Tris-HCl, 100 mM KCl, and 10 mM $MgCl_2$ at pH 7.4.

The K^+ ions in the sensor buffer trigger G-quadruplex formation, and the typical signature of a G-quadruplex was readily observed in the unmodified RNA aptamer (UbSP) as shown in Figure 2C (red trace). The CD spectrum of the UbSP aptamer also revealed a positive band at 270 nm and smaller negative peaks at 238 and 215 nm, indicating the formation of a mixture of the hybrid, parallel/antiparallel G-quadruplex, and is consistent with what has been reported in the literature.[60,61] Interestingly, a similar CD result was obtained for the modified BAbSP RNA aptamer as well (Figure 2C, blue trace), in which up to 10 uridine residues could have been potentially replaced by the PGN, barbitudine. The overall secondary structure seemed to be in reasonable agreement with the control UbSP aptamer. However, the ellipticity values on the *Y*-axis seem to indicate a marginally altered effect that could be coming from the presence of the multiple barbitudine residues in the BAbSP aptamer. This could be due to the inability of the barbitudine residues to stack and base pair effectively with cognate and neighboring nucleobases.[62] To assess this, UV-thermal melting analysis of control UbSP and modified BAbSP aptamers was performed. Five μM unmodified (UbSP) and 5 μM modified (BAbSP) RNAs were prepared in a sensor buffer following the exact procedure for the preparation of samples for CD analysis. The melting studies show that the modified BAbSP aptamer has a lower melting temperature (58 °C ± 0.5 °C) when compared to the control unmodified UbSP aptamer (70 °C ± 0.5 °C, Figure S19). The thermal melting profile emphasized that after incorporation into the aptamer RNA, the 6-oxo group of the barbitudine was affecting the H-bonding. Additionally, poor stacking of barbitudine[62] would have further resulted in lowering the melting temperature for BAbSP when compared to the control UbSP aptamer. These experiments also indicate that the H-bonding and stacking ability of the heterocycle could have been a significant selection pressure on pre-RNAs to undergo further evolution to result in more robust functional polymers.

Biophysical Analysis of the Modified Baby Spinach RNA Aptamers using Fluorescence Spectroscopy. To act as a Darwinian ancestor, the PGN containing aptamer must demonstrate the ability to fold into the requisite secondary structure that would result in emergent functions like ligand binding in the case of an aptamer, or facilitate catalysis, etc. To test this possibility, we first designed and synthesized the aptamer RNA that contained the noncanonical PGN, BaTP. CD analysis confirmed the ability of this modified RNA aptamer (BAbSP) to adopt a functionally active secondary structure. To further affirm whether the BAbSP aptamer was indeed in the "active" form, we performed fluorescence analysis of BAbSP in the presence of the ligand, DFHBI (3,5-difluoro-4-hydroxyphenyl)methylene]-3,5-dihydro-2,3-dimethyl-4H-imidazole-4-one). A fluorescence enhancement after binding to the DFHBI ligand would help confirm the ability of the barbitudine-modified RNA, BAbSP, to adopt the correct secondary structure. Importantly, it will establish its ability to acquire "function," which in this case is that of an aptameric function of ligand binding.

Given the aforesaid, the binding of both the control and modified versions of the bSP RNA aptamers, to the nonfluorescent dye DFHBI, was studied as this ligand results in fluorescence signal enhancement on binding to the target aptamer. Samples of 5 μM unmodified UbSP and modified BAbSP aptamer RNAs, respectively, were prepared in a sensor buffer using the protocol described for sample preparation for CD analysis. To these samples was added 10 μM DFHBI ligand, and the samples were incubated in the dark for 2h. All of the respective control samples were also prepared in a similar manner. After the incubation period, all the samples were analyzed using fluorescence spectroscopy. RNA samples (UbSP and BAbSP) in the absence of the ligand DFHBI (L) did not show any detectable fluorescence (Figure 3C, red and blue traces, respectively). Further, the DFHBI-alone control showed very weak fluorescence in the absence of the target aptamer RNA (Figure 3C, black trace). However, both aptamer RNAs (UbSP and BAbSP) showed a significant increase in fluorescence intensity when incubated with the DFHBI ligand. The unmodified UbSP aptamer displayed a 5-

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

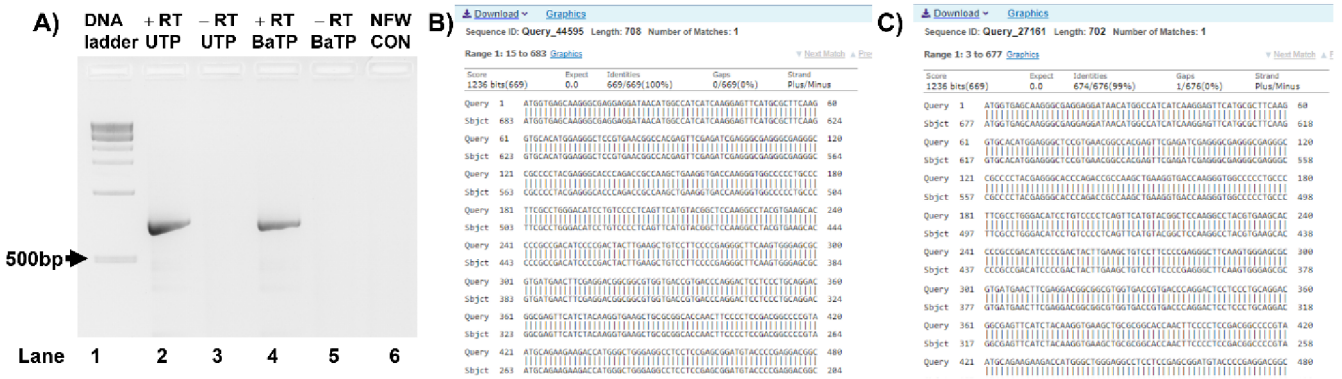


Figure 4. RT-PCR and sequence alignment results for unmodified and modified mCherry RNA. (A) Gel picture depicting PCR amplicons obtained from the various cDNA samples. PCR reactions were performed with cDNA that was generated after the successful reverse transcription of control RNA containing UTP and modified RNA containing BaTP (+RT samples), respectively. DNase-treated RNA was used as the negative reaction control (−RT). Representative sequence alignment (BLAST) of PCR amplified DNA products obtained from cDNA, which was in turn obtained by reverse transcribing either (B) unmodified RNA (containing uridine) or (C) modified RNA (containing barbitudine). The PCR product sequencing data showed 99% sequence identities with the original mCherry template DNA sequence that was used for the incorporation of UTP or BaTP into the RNA (using IVT).

fold increase (Figure 3C, pink trace) in the fluorescence intensity, while the modified BAbSP aptamer showed a nearly 4-fold increase (Figure 3C, green trace). Overall, the fluorescence analysis indicated that the function of ligand binding to the aptamer RNAs was clearly happening, but the efficiency of ligand binding to the modified aptamer and the underlying mechanism is to be ascertained.

Information Transfer from Barbitudine Containing RNA to DNA through Reverse Transcription and PCR Amplification. Given the success with the T7 polymerase incorporation of BaTP, we set out to evaluate the propensity of other extant enzymes to use BaTP effectively. Since PCR amplification and sequencing would allow a measure of the fidelity of canonical enzymes for tolerating the incorporation of a PGN like BaTP, evaluating enzymes relevant to this process was the logical next step. Toward this, UbSP and BAbSP RNAs were subjected to reverse transcription using Superscript-III reverse transcriptase. The cDNA obtained from these two separate reverse transcription reactions were PCR amplified using high-fidelity Pfu DNA polymerase in the presence of appropriate forward and reverse primers (Materials and Methods section). The products obtained from the PCR reactions were then processed for Sanger sequencing. However, due to the product DNAs' (small) length, there was technical difficulty in sequencing them. An alternate strategy was to monitor the incorporation of BaTP into a longer RNA, which could then be subjected to RT-PCR and sequencing analysis.[63] Toward this, an ∼700 nt long mCherry DNA was transcribed using T7 RNA polymerase by an *in vitro* transcription reaction, using a DNA template in the presence of UTP (control reaction) or BaTP (modified reaction) plus ATP, CTP and GTP nucleotides. To confirm the incorporation of BaTP, we performed the enzymatic digestion of the resultant mCherry mRNA and subjected this digest to HPLC analysis. The HPLC profile showed the incorporation of barbitudine into the modified mCherry mRNA (Figure S20).

Subsequent to this, both the control and the modified RNA transcripts were subjected to RNase-free DNase treatment for 1h at 37 °C, followed by purification using LiCl precipitation method. To generate the complementary DNA (cDNA), 1 μg of the respective RNA samples were subjected to reverse transcription using Superscript III reverse transcriptase and random hexamer reverse primers. Further, the cDNAs generated from the respective RNAs were subjected to PCR amplification using a high-fidelity Pfu DNA polymerase in the presence of appropriate forward and reverse primers. We used the following controls to confirm the successful reverse transcription of the RNA transcripts; i.e., the minus RT control (reaction without the RT enzyme) and the NTC control (a nontemplate control). PCR products obtained after the amplification reaction were directly resolved on 1.5% agarose gel to confirm cDNA synthesis (Figure 4A). Notably, all the positive (or test) reverse transcription reactions (+RT samples) showed the amplicon band of the same length (Figure 4A, **lanes 2 and 4**). Importantly, the absence of a band from the negative "minus RT" control reactions further confirmed that there was indeed a successful information transfer happening from the control U and the modified barbitudine containing RNA transcripts into the cDNA, respectively. This confirmed the incorporation of A against BA in the modified mCherry RNA transcripts. Further, to evaluate the fidelity of the transcription and reverse transcription reactions, PCR amplicons obtained from the amplification of the respective cDNA products were validated by Sanger sequencing (Figures S21−S24).

The sequence analysis of the PCR products showed 99% sequence identity with the mCherry template DNA sequence that was originally used for the incorporation of UTP or BaTP into the RNAs during the transcription reaction step. Importantly, sequence blasts also clearly showed that no misincorporation was observed while incorporating thymidine against the adenosine residues that were present in the resultant cDNA template (Figures S22−S24). Taken together, these results confirmed that BaTP was efficiently incorporated into the RNA transcripts (against cognate deoxyadenosine), which was then effectively recognized and copied by reverse transcriptase to produce the corresponding cDNA by incorporating T against the barbitudine in the pre-RNA template). This ability of BaTP to be recognized and used in a series of steps, namely, transcription, reverse transcription, and amplification, would have been beneficial for faithful information transfer in a pre-RNA World and its subsequent transition to the RNA−DNA−protein world.

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

CONCLUSIONS

One of the fundamental hallmarks of life is its tremendous capacity for evolution, which manifests at various levels. Successful attempts to expand the genetic code of extant life to create artificial or semisynthetic life confirm that evolution can be tinkered with by externally facilitating changes (by applying relevant "selection pressures"), which result in consequences at the molecular level. The effective proliferation of life requires molecules that can store and propagate information and perform catalysis. It is hypothesized that RNA informational molecules that acquired functional capabilities, like in aptamers and ribozymes, would have been selected for further evolution during the chemical origins of life.[64] Nonetheless, the emergence of a molecule like RNA readily from a heterogeneous prebiotic soup is a nontrivial process even under the best of circumstances.[65] Given this, the systematic understanding of how primitive informational polymers of a pre-RNA World emerged and functioned is fundamental to understanding how biotic life could have evolved from abiotic life.

In this backdrop, we report the synthesis and enzymatic incorporation of a prebiotically pertinent genetic nucleotide BaTP, using prokaryotic proteinaceous T7 RNA polymerase. To the best of our knowledge, such enzymatic incorporation of BaTP has been reported for the first time. We observed that BaTP can be effectively incorporated by this old-world RNA polymerase to result in a full-length RNA product. Additionally, BaTP also competes with its canonical counterpart UTP during the transcription reaction. Understandably, the preference of T7 polymerase, qualitatively speaking, seems to be more for the canonical nucleotide than the noncanonical one. This could be an indication of the early selection pressures that could have been acting on such PGNs. Nevertheless, more systematic quantitative studies are needed to affirm this. Further, we also demonstrated the ability of these putative pre-RNA aptamers to adopt secondary structures by using CD analysis. This result suggests a H-bonding pattern that could be shared between uridine and barbitudine, as the CD profiles of the unmodified and modified RNA aptamers did not indicate any drastic perturbation to the secondary structure. Finally, these results were further confirmed using fluorescence analysis wherein we showed that the aptamer obtained by incorporation of a PGN (BaTP) did not lead to the hampering of its function, which is that of binding to its ligand. *Au contraire*, it showed effective binding to its ligand (DFHBI) and there was fluorescence enhancement comparable to that of the unmodified UbSP aptamer.

Collectively, these results indicate that the prebiotic genetic nucleotide barbitudine has the potential to act as an early Darwinian ancestor of modern pyrimidine nucleotides. Further, we have systematically undertaken experiments showing the ability for information transfer from barbitudine and the fidelity associated with the related events. We also showed evidence for the barbitudine containing RNA to act as a template for the synthesis of complementary DNA in the presence of a high-fidelity reverse transcriptase. This cDNA could also be further amplified using a PCR reaction to yield a full-length DNA product. These very compelling observations highlight the potential of pre-RNAs containing barbitudine to be suitable for processes like transcription and reverse transcription, two processes integral to the central dogma of extant biology. We are currently exploring other such PGNs and the ability of mRNAs that have these PGNs to potentially undergo protein synthesis, with multiple ramifications including for interesting *in vitro* and *in cellulo* studies.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c09949.

Supporting Information: ^{1}H, ^{13}C, and ^{31}P NMR and HRMS spectra of all the compounds, HPLC traces, HRMS of aptamer RNA digests, denaturing gel image of the aptamer RNAs, UV−Tm melting spectra, HPLC of enzymatic digestion of mCherry RNA, sequencing data (PDF)

AUTHOR INFORMATION

Corresponding Authors

Anupam A. Sawant − *Department of Biology, Indian Institute of Science Education and Research (IISER) Pune, Pune, Maharashtra 411008, India*; Email: anupam.sawant@iiserpune.ac.in

Sudha Rajamani − *Department of Biology, Indian Institute of Science Education and Research (IISER) Pune, Pune, Maharashtra 411008, India*; orcid.org/0000-0002-6030-495X; Email: srajamani@iiserpune.ac.in

Authors

Sneha Tripathi − *Department of Biology, Indian Institute of Science Education and Research (IISER) Pune, Pune, Maharashtra 411008, India*

Sanjeev Galande − *Department of Biology, Indian Institute of Science Education and Research (IISER) Pune, Pune, Maharashtra 411008, India; Center of Excellence in Epigenetics, Department of Life Sciences, School of Natural Sciences, Shiv Nadar Institution of Eminence, Uttar Pradesh 201314, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c09949

Author Contributions

A.A.S. and S.R. conceived the project and designed the experiments. A.A.S. performed the chemical synthesis of the nucleoside triphosphate (BaTP) and performed all the biophysical experiments and analysis, the enzymatic digestion assays and related experiments. S.T. performed in vitro transcription reactions, RT-PCR reactions, and gel analysis with inputs from S.G. A.A.S. and S.R. wrote the manuscript with critical inputs from S.T. and S.G. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.A.S thanks DBT-RA, DBT, Government of India for funding support. S.R. thanks SERB, DST [CRG/2021/001851], and DBT, Government of India [BT/PR19201/BRB/10/1532/2016], for funding. S.G. is a recipient of the JC Bose Fellowship (JCB/2019/000013) from the Science and Engineering Research Board, Government of India. S.R. and A.A.S. express their special gratitude to Dr. Srivatsan's and Dr. Hotha's laboratories for help with certain reagents and for providing laboratory access for synthesis work.

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

REFERENCES

(1) Gilbert, W. The RNA world. *Nature* **1986**, *319*, 618.
(2) Joyce, G. F. RNA evolution and the origins of life. *Nature* **1989**, *338*, 217−224.
(3) Walker, S. I.; Packard, N.; Cody, G. D. Re-conceptualizing the origins of life. *Philos. Trans. R. Soc. A* **2017**, *375*, 20160337.
(4) Jortner, J. Conditions for the emergence of life on the early Earth: summary and reflections. *Philos. Trans R Soc. Lond B Biol. Sci.* **2006**, *361*, 1877−1891.
(5) Woese, C. The universal ancestor. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 6854−6859.
(6) Koonin, E. V. Comparative genomics, minimal gene sets, and the last universal common ancestor. *Nature Rev. Microbiol.* **2003**, *1*, 127−136.
(7) Miller, S. L.; Urey, H. C. Organic compound synthesis on the primitive earth. *Science* **1959**, *130*, 245−251.
(8) Salván, C. M.; Bouza, M.; Fialho, D. M.; Burcar, B. T.; Fernández, F. M.; Hud, N. V. Prebiotic Origin of Pre-RNA Building Blocks in a Urea "Warm Little Pond" Scenario. *ChemBiochem* **2020**, *21*, 3504−3510.
(9) Grosch, M.; Stiebritz, M. T.; Bolney, R.; Winkler, M.; Jückstock, E.; Busch, H.; Peters, S.; Siegle, A. F.; van Slageren, J.; Ribbe, M.; et al.etal Mackinawite-Supported Reduction of C1 Substrates into Prebiotically Relevant Precursors. *ChemSystemschem* **2022**, *4*, No. e202200010.
(10) Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **2015**, *7*, 301−307.
(11) Morowitz, H. J.; Kostelnik, J. D.; Yang, J.; Cody, G. D. The origin of intermediary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 7704−7708.
(12) Stewart, J. E. The Origins of Life: The Managed-Metabolism Hypothesis. *Found Sci* **2019**, *24*, 171−195.
(13) Andras, P.; Andras, C. The origins of life − the 'protein interaction world' hypothesis: protein interactions were the first form of self-reproducing life and nucleic acids evolved later as memory molecules. *Med. Hypotheses* **2005**, *64*, 678−688.
(14) Müller, F.; Escobar, L.; Xu, F.; Węgrzyn, E.; Nainytė, M.; Amatov, T.; Chan, C.-Y.; Pichler, A.; Carell, T. A prebiotically plausible scenario of an RNA−peptide world. *Nature* **2022**, *605*, 279−284.
(15) Segre, D.; Ben-Eli, D.; Deamer, D. W.; Lancet, D. The Lipid World. Origins Life. *Evol. Biospheres* **2001**, *31*, 119−145.
(16) Deamer, D.; Dworkin, J. P.; Sandford, S. A.; Bernstein, M. P.; Allamandola, L. J. The first cell membranes. *Astrobiology* **2002**, *2*, 371−381.
(17) Joyce, G. F. The antiquity of RNA-based evolution. *Nature* **2002**, *418*, 214−221.
(18) Higgs, P.; Lehman, N. The RNA World: molecular cooperation at the origins of life. *Nat. Rev. Genet.* **2015**, *16*, 7−17.
(19) Pressman, A.; Blanco, C.; Chen, I. A. The RNA World as a Model System to Study the Origin of Life. *Curr. Biol.* **2015**, *25*, R953−R963.
(20) Pearce, B. K. D.; Pudritz, R. E.; Semenov, D. A.; Henning, T. K. Origin of the RNA world: The fate of nucleobases in warm little ponds. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 11327−11332.
(21) Yarus, M. Getting past the RNA world: The initial Darwinian ancestor. *Cold Spring Harb Perspect Biol.* **2011**, *3*, a003590.
(22) Cafferty, B. J.; Hud, N. V. Abiotic synthesis of RNA in water: a common goal of prebiotic chemistry and bottom-up synthetic biology. *Curr. Opin. Chem. Biol.* **2014**, *22*, 146−157.
(23) Kruse, F. M.; Teichert, J. S.; Trapp, O. Prebiotic Nucleoside Synthesis: The Selectivity of Simplicity. *Chem.-Eur. J.* **2020**, *26*, 14776−14790.
(24) Yadav, M.; Kumar, R.; Krishnamurthy, R. Chemistry of Abiotic Nucleotide Synthesis. *Chem. Rev.* **2020**, *120*, 4766−4805.
(25) Yi, J.; Kaur, H.; Kazöne, W.; Rauscher, S. A.; Gravillier, L. A.; Muchowska, K. B.; Moran, J. A Nonenzymatic Analog of Pyrimidine Nucleobase Biosynthesis. *Angew. Chem., Int. Ed. Engl.* **2022**, *61*, No. e202117211.
(26) Saladino, R.; Šponer, J. E.; Šponer, J.; Di Mauro, E. Rewarming the Primordial Soup: Revisitations and Rediscoveries in Prebiotic Chemistry. *ChemBiochem* **2018**, *19*, 22−25.
(27) Callahan, M. P.; Smith, K. E.; Cleaves, H. J., II; Ruzicka, J.; Stern, J. C.; Glavin, D. P.; House, C. H.; Dworkin, J. P. Carbonaceous meteorites contain a wide range of extraterrestrial nucleobases. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13995−13998.
(28) Kolb, V. M.; Dworkin, J. P.; Miller, S. L. Alternative bases in the RNA world: the prebiotic synthesis of urazole and its ribosides. *J. Mol. Evol.* **1994**, *38*, 549−557.
(29) Joyce, G. F.; Schwartz, A. W.; Miller, S. L.; Orgel, L. E. The case for an ancestral genetic system involving simple analogs of the nucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 4398−4402.
(30) Hud, N. V.; Cafferty, B. J.; Krishnamurthy, R.; Williams, L. D. The origin of RNA and'My Grandfather's Axe'. *Chem. Biol.* **2013**, *20*, 466−474.
(31) Hud, N. V. Searching for lost nucleotides of the pre-RNA World with a self-refining model of early Earth. *Nat. Commun.* **2018**, *9*, 5171−5175.
(32) Bean, H. D.; Sheng, Y.; Collins, J. P.; Anet, F. A. L.; Leszczynski, J.; Hud, N. V. Formation of a β-pyrimidine nucleoside by a free pyrimidine base and ribose in a plausible prebiotic reaction. *J. Am. Chem. Soc.* **2007**, *129*, 9556−9557.
(33) Chen, M. C.; Cafferty, B. J.; Mamajanov, I.; Gállego, I.; Khanam, J.; Krishnamurthy, R.; Hud, N. V. Spontaneous prebiotic formation of a β-ribofuranoside that self-assembles with a complementary heterocycle. *J. Am. Chem. Soc.* **2014**, *136*, 5640−5646.
(34) Becker, S.; Feldmann, J.; Wiedemann, S.; Okamura, H.; Schneider, C.; Iwan, K.; Crisp, A.; Rossa, M.; Amatov, T.; Carell, T. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. *Science* **2019**, *366*, 76−82.
(35) Xu, F.; Crisp, A.; Schinkel, T.; Dubini, R. C. A.; Hübner, S.; Becker, S.; Schelter, F.; Rovó, P.; Carell, T. Isoxazole Nucleosides as Building Blocks for a Plausible Proto-RNA. *Angew. Chem., Int. Ed. Engl.* **2022**, *61*, No. e202211945.
(36) Mungi, C. V.; Singh, S. K.; Chugh, J.; Rajamani, S. Synthesis of barbituric acid containing nucleotides and their implications for the origin of primitive informational polymers. *Phys. Chem. Chem. Phys* **2016**, *18*, 20144−20152.
(37) Cafferty, B. J.; Hud, N. V. Was a pyrimidine-pyrimidine base pair the ancestor of Watson-Crick base pairs? Insights from a systematic approach to the origin of RNA. *Israel J. Chem.* **2015**, *55*, 891−905.
(38) Ellington, A. D.; Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818−822.
(39) Wilson, D. S.; Szostak, J. W. In vitro selection of functional nucleic acids. *Annu. Rev. Biochem.* **1999**, *68*, 611−647.
(40) Carothers, J. M.; Oestreich, S. C.; Szostak, J. W. Aptamers selected for higher-affinity binding are not more specific for the target ligand. *J. Am. Chem. Soc.* **2006**, *128*, 7929−7937.
(41) Romesberg, F. E. Discovery, implications and initial use of semi-synthetic organisms with an expanded genetic alphabet/code. *Philos. Trans. R. Soc. B* **2023**, *378*, 20220030.
(42) Manandhar, M.; Chun, E.; Romesberg, E. F. Genetic Code Expansion: Inception, Development, Commercialization. *J. Am. Chem. Soc.* **2021**, *143*, 4859−4878.
(43) Malyshev, D.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Corrêa, I. R.; Romesberg, F. E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **2014**, *509*, 385−388.
(44) Zhang, Y.; Ptacin, J.; Fischer, E.; Aerni, H. R.; Caffaro, C. E.; San Jose, K.; Feldman, A. W.; Turner, C. R.; Romesberg, F. E. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **2017**, *551*, 644−647.
(45) Hoshika, S.; Leal, N. A.; Kim, M.-J.; Kim, M. S.; Karalkar, N. B.; Kim, H. J.; Bates, A. M.; Watkins Jr, N. E.; SantaLucia, H. A.; Meyer, A. J.; Dasgupta, S.; Piccirilli, J. A.; Ellington, A. D.; SantaLucia Jr, J.;

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082

Georgiadis, M. M.; Benner, S. A. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science* **2019**, *363*, 884−887.
(46) Henry, A. A.; Romesberg, F. E. Beyond A, C, G, and T: augmenting nature's alphabet. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727−733.
(47) Anderson, B. A.; Fauché, K.; Karunakaran, S. C.; Yerabolu, J. R.; Hud, N. V.; Krishnamurthy, R. The Unexpected Base-Pairing Behavior of Cyanuric Acid in RNA and Ribose versus Cyanuric Acid Induced Helicene Assembly of Nucleic Acids: Implications for the Pre-RNA Paradigm. *Chemistry* **2021**, *27*, 4033−4042.
(48) Cafferty, B. J.; Fialho, D. M.; Khanam, J.; Krishnamurthy, R.; Hud, N. V. Spontaneous formation and base pairing of plausible prebiotic nucleotides in water. *Nat. Commun.* **2016**, *7*, 11328.
(49) Mungi, C. V.; Rajamani, S. Characterization of RNA-Like Oligomers from Lipid-Assisted Nonenzymatic Synthesis: Implications for Origin of Informational Molecules on Early Earth. *Life* **2015**, *5*, 65−84.
(50) Charette, M.; Gray, M. W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **2000**, *49*, 341−351.
(51) Lin, H.; Jiménez, E. I.; Arriola, J. T.; Müller, U. F.; Krishnamurthy, R. Concurrent Prebiotic Formation of Nucleoside-Amidophosphates and Nucleoside-Triphosphates Potentiates Transition from Abiotic to Biotic Polymerization. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202113625.
(52) Whitaker, D.; Powner, M. W. Prebiotic synthesis and triphosphorylation of 3′-amino-TNA nucleosides. *Nat. Chem.* **2022**, *14*, 766−774.
(53) El-Tayeb, A.; Qi, A.; Müller, C. E. Synthesis and Structure−Activity Relationships of Uracil Nucleotide Derivatives and Analogues as Agonists at Human P2Y 2 , P2Y 4 , and P2Y 6 Receptors. *J. Med. Chem.* **2006**, *49*, 7076−7087.
(54) Dietzsch, J.; Bialas, D.; Bandorf, J.; Würthner, F.; Höbartner, C. Tuning Exciton Coupling of Merocyanine Nucleoside Dimers by RNA, DNA and GNA Double Helix Conformations. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202116783.
(55) Niedballa, U.; Vorbrüggen, H. A general synthesis of pyrimidine nucleosides. *Angew. Chem., Int. Ed. Engl.* **1970**, *9*, 461−462.
(56) Milligan, J. F.; Groebe, D. R.; Witherell, G. W.; Uhlenbeck, O. C. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **1987**, *15*, 8783−8798.
(57) Milligan, J. F.; Uhlenbeck, O. C. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **1989**, *180*, 51−62.
(58) Okuda, M.; Fourmy, D.; Yoshizawa, S. Use of Baby Spinach and Broccoli for imaging of structured cellular RNAs. *Nucleic Acids Res.* **2017**, *45*, 1404−1415.
(59) Dagar, S.; Sarkar, S.; Rajamani, S. Nonenzymatic Template-Directed Primer Extension Using 2′-3′ Cyclic Nucleotides Under Wet-Dry Cycles. *Orig. Life Evol. Biosph.* **2023**, *53*, 43−60.
(60) Dao, N. T.; Haselsberger, R.; Khuc, M. T.; Phan, A. T.; Voityuk, A. A.; Michel-Beyerle, M. E. Photophysics of DFHBI bound to RNA aptamer Baby Spinach. *Sci. Rep.* **2021**, *11*, 7356.
(61) DasGupta, S.; Shelke, S. A.; Li, N.-S.; Piccirilli, J. A. Spinach RNA aptamer detects lead (II) with high selectivity. *Chem. Commun.* **2015**, *51*, 9034−9037.
(62) Salameh, H.; Afri, M.; Gottlieb, H. E.; Fischer, B. *β*-Cyanuryl Ribose, *β*-Barbituryl Ribose, and 6-Azauridine as Uridine Mimetics. *ACS Omega* **2020**, *5*, 31314−31322.
(63) Sawant, A. A.; Tanpure, A. A.; Mukherjee, P. P.; Athavale, S.; Kelkar, A.; Galande, S.; Srivatsan, S. G. A versatile toolbox for posttranscriptional chemical labeling and imaging of RNA. *Nucleic Acids Res.* **2016**, *44*, No. e16.
(64) Song, E. Y.; Jiménez, E. I.; Lin, H.; Le Vay, K.; Krishnamurthy, R.; Mutschler, H. Prebiotically Plausible RNA Activation Compatible with Ribozyme-Catalyzed Ligation Angew. *Chem., Int. Ed.* **2021**, *60*, 2952−2957.
(65) Schneider, C.; Becker, S.; Okamura, H.; Crisp, A.; Amatov, T.; Stadlmeier, M.; Carell, T. Noncanonical RNA Nucleosides as Molecular Fossils of an Early Earth-Generation by Prebiotic Methylations and Carbamoylations. *Angew. Chem., Int. Ed. Engl.* **2018**, *20*, 5943−5946.

https://doi.org/10.1021/acsomega.3c09949
ACS Omega 2024, 9, 18072−18082