N2#P3# Phosphoramidate Glycerol Nucleic Acid as a Potential Alternative Genetic System

Jesse J. Chen, Xin Cai, and Jack W. Szostak

J. Am. Chem. Soc., 2009, 131 (6), 2119-2121 • DOI: 10.1021/ja809069b • Publication Date (Web): 23 January 2009

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML
N2′→P3′ Phosphoramidate Glycerol Nucleic Acid as a Potential Alternative Genetic System

Jesse J. Chen, Xin Cai, and Jack W. Szostak*

Howard Hughes Medical Institute, and Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114

Received November 19, 2008; E-mail: szostak@molfbio.mgh.harvard.edu

Published on Web 01/23/2009

Glycerol nucleic acid (GNA), which is based upon an acyclic, three-carbon backbone (Figure 1A), is a promising starting point for the construction of an artificial genetic system owing to its ability to form stable, antiparallel duplexes.1,2 As proposed by Orgel,3 we imagine that a nonenzymatic, template-directed polymerization process mediated replication during the evolution of early genetic systems,4,5 in place of enzymatic replication as in the modern genetic system. To recapitulate this process, our group has focused on the chemistry of polymerization facilitated by an imidazole-activated phosphate group and an amino group as the nucleophile (Figure 1B and 1C, highlighted in red).5 This approach requires the newly synthesized strand with N→P phosphoramidate linkages to form stable, Watson–Crick base pairs with the template strand. The base-pairing properties of the GNA analogue with N2′-amino GNA dinucleotides (Figure 1C). These results suggest that npGNA is a potential candidate for a self-replicating system based on phosphoramidate linkages.

Figure 1. GNA-based genetic system. (A) Structures of GNA and npGNA. (B and C) Chemistry of templated synthesis of npGNA using imidazole-activated monomer (B) or dimer (C). The nucleophile and leaving group involved in each transformation are highlighted in red.

npGNA oligomers were synthesized via a solid-phase, oxidative amination coupling method1,6 using 2′-amino glycerol activated monomer (B) or dimer (C). The nucleophile and leaving group are highlighted in red.

Amination Coupling

Scheme 1. Synthesis of 2′-Amino Glycerol Nucleoside Monomers

Scheme 2. Synthesis of npGNA Using Solid-Phase, Oxidative Amination Coupling

For 1t→4t, B = thymine; for 1c→4c, B = N4-benzoylcytosine; for 1a→4a, B = N4-benzoyladenine; for 1g→4g, B = N4-isobutyrylguanine. For 3t, R1 = O, R2 = CH2; for 3c, R1 = NBz, R2 = H. Conditions and yields: (a) DAST, pyridine/CH2Cl2; 3t, not isolated; 3c, 54%. (b) NaN3, 1:1 DMP/DMF, 110 °C; 4t, 42%; 4c, 46%. (c) (i) MesCl, Et3N/CH2Cl2; (ii) NaH, 1:1 DMP/DMF, 100 °C; 4a, 71%; 4g, 96%. (d) For 4t and 4g, H2, Pd/C, 78% and 77%, respectively; for 4c and 4a, H2, 15% Et3N in Pyridine, 0 °C, 83% and 91%, respectively.

Conditions: (a) 3% dichloroacetic acid in CH2Cl2; (b) 0.4 M 2-cyanoethyl-N,N-diisopropylchlorophosphite/0.4 M diisopropylethylamine in CH2Cl2; (c) 0.4 M tetrazole in 9:1 acetonitrile/water; (d) 0.2 M Ht, 1c, 1a, or 1g and 0.2 M triethylamine in 1:1 CCl4/acetonitrile. Functional groups involved in each transformation are highlighted in red.

thymidine at the 2′-terminus in all npGNA sequences synthesized in this study (Table S1).

This method was validated by the synthesis of dinucleotide model compounds (5t→5g, Figure S1). Only one diastereomer was observed for all 4 dinucleotides in 1H and 31P NMR spectra, indicating a clean SN2 attack by azide leading to 4t→4g. In addition, the dinucleotides had 31P chemical shifts of ~7 to 8 ppm (Supporting Information, Figure S1) and underwent acid-catalyzed decomposition, both characteristic of a phosphoramidate linkage. The coupling efficiency of each cycle is ~40% for 1a and 70–80% for 1t, 1c, and 1g. The decreased coupling efficiency compared with phosphoramidate DNA synthesis is thought to be related to the decreased nucleophilicity of the 2′-NH2 due to the presence of...
sequences (Table 1). The duplex formation was also observed in several other npGNA of each strand was established by a Job plot (Figure S3). Stable npGNA formed a stable duplex (Figure 2). The 1:1 stoichiometry UV and CD experiments were performed on gel- or HPLC-purified and characterized by MALDI-TOF MS (Table S1). All subsequent mide gel electrophoresis or by anion-exchange HPLC (Figure S2) were quantitatively recovered after each coupling step. Crude isolate several npGNA oligomers (up to 10-mer) in amounts sufficient for base-pairing studies. The unreacted monomers (Figure 2) and CD (Figure S5) measured by CD experiments were consistent with those obtained of the same sequence (Table 1). Formation of duplex npGNA was confirmed by circular dichroism (CD) spectroscopy, which revealed a temperature-dependent conformational change upon denaturation of npGNA duplexes (Figures 3 and S4). The transition temperatures measured by CD experiments were consistent with those obtained in UV thermal denaturation studies (Table 1).

Thermal denaturation (Figure 2) and CD (Figure S5–S7) studies also revealed stable heteroduplex formation between GNA and npGNA. Comparison of the $T_m$ values in the same sequence context suggests that thermostability increases in the order of npGNA, npGNA:GNA, and GNA:GNA duplexes (Table 1). Interestingly, the reverse trend of thermostability was reported for N3′−P5′ phosphoramidate DNA (npDNA), which increases in the order of DNA, npDNA:DNA, and npDNA duplexes. Previous X-ray structural studies and molecular dynamics simulations revealed that npDNA adopts an A-form conformation and its backbone is more rigid than that of DNA. We found that GNA, GNA:npGNA, and npGNA duplexes share similar CD features (Figure S5), suggesting that all three share a similar conformation in solution.

We then investigated the template-dependent synthesis of npGNA oligomers using 3′-imidazole-activated precursors (Figure 1B). A photolytic protecting group (6-nitroveratryl carbamate or Nvoc) was used to protect the 2′-NH$_2$ before activation of the 3′-phosphate group with carbonyl diimidazole (10a, Scheme 3A). After photolysis, however, we found that the imidazolide 11a underwent rapid intramolecular cyclization at pH $<13$ to yield 12a (Scheme 3A). The rate of cyclization was too fast to be accurately measured by either NMR or HPLC analysis under the conditions for templated polymerization (at 4°C, pH 8.4). However, we found that an activated dinucleotide (14a), which can be generated by photodeprotection of its precursor (13a, Scheme 3B), was much more stable than 12a with a half-life of 53.7 h at 4°C and pH 8.4 as measured by HPLC (Figure S8). The breakdown products of 14a were cyclic and linear dinucleotides with a ratio of $\sim$3:1 (Figure S8).

We envisioned that 13a could be used directly in templated polymerization reactions by generating 14a in situ via photolysis. Because GNA forms a stable duplex with npGNA and can be conveniently synthesized in a large quantity, GNA-(T)$_{10}$ was used as the template for studying polymerization of 14a. As shown in Figure 4A, we observed polymerization of 14a by MALDI-TOF MS analysis. The full-length product (10-mer) was detected together with shorter oligomers after photolysis (Figure 4A). In the absence of the GNA template, only the 4-mer and a small amount of 6-mer products were observed with no full-length or 8-mer product (Figure 4B), suggesting that the template is necessary for efficient polymerization of 14a. We also investigated incorporation of 14a into a primer-template complex. Instead of using a GNA or npGNA
In summary, we have demonstrated that npGNA can form stable homo- and heteroduplexes with a conformation similar to that of GNA. In addition, N2'→P3' linkages can be formed by template-dependent polymerization, which suggests that sequence information transfer can be achieved nonenzymatically in a genetic system based on GNA oligonucleotides.

**Acknowledgment.** J.J.C. was supported by a Fellowship from the Harvard Origins Initiative. This work was supported by NSF grant CHE 0434507 to J.W.S. J.W.S is an Investigator of the Howard Hughes Medical Institute. The authors thank Dr. A. Ricardo, Dr. M. Krishnamurthy, Dr. D. Treco, and Dr. B. Heuberger for helpful discussions.

**Supporting Information Available:** Procedures for synthesis and characterization of 1t-1g, 5t-5g, and 6a-14a, and solid-phase synthesis of npGNA, HPLC profile of the npGNA-(T)10 crude product, MALDI-TOF analysis of GNA and npGNA oligonucleotides in this study, UV melting curves, and additional CD spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

### References


**Figure 5.** Template-dependent polymerization of 14a analyzed by primer extension: (lane 1) primer only as control; (lane 2) without template; (lane 3) with template; (lane 4) with control template. The reaction mixture of 10 µL contained 1 µM 5'-[32P]-primer/template, 2 mM 13a, 20 mM MgCl2, 0.2 M NaCl, 0.1 M HEPBS, pH 8.4. The reaction was initiated by photolysis at 4 °C for 6 h and, after another 12 h incubation at 4 °C, was analyzed by 20% denaturing polyacrylamide gel electrophoresis.

**Figure 4.** Template-dependent polymerization of the activated dinucleotide 14a on T-3'-TTT TTT TTT T-2'T analyzed by MALDI-TOF MS. (The underlined denotes the GNA 10-mer sequence): (A) with the template; (B) without the template. In B, the GNA template was cosootted with the sample as an internal control. The polymerization products were identified by comparing the signals to their expected masses. [M + H]+ for 4-mer: calcd, 1101.733; obsd, 1100.01. For 6-mer: calcd, 1642.349; obsd, 1640.67. For 8-mer: calcd, 2184.443; obsd, 2182.59. For 10-mer: calcd, 2725.798; obsd, 2723.68. For the template: calcd, 3168.994; obsd, 3167.08. (* denotes the n-1 synthetic byproduct of the template).