

HETEROCYCLES, Vol. 82, No. 2, 2011, pp. 1477 - 1488. © The Japan Institute of Heterocyclic Chemistry  
Received, 28th July, 2010, Accepted, 13th September, 2010, Published online, 14th September, 2010  
DOI: 10.3987/COM-10-S(E)99

## STEP-GROWTH CONTROL IN TEMPLATE-DIRECTED POLYMERIZATION\*\*

Xiaoyu Li,<sup>\*§</sup> Andres F. Hernandez,<sup>o‡</sup> Martha A. Grover,<sup>o‡</sup> Nicholas V. Hud,<sup>\*o</sup> and David G. Lynn<sup>\*o§</sup>

\* Center for Fundamental and Applied Molecular Evolution

o Center for Chemical Evolution

‡ School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

• School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA

§ Departments of Chemistry and Biology, Emory University, Atlanta, GA 30322, USA

Contacts:

MAG: martha.grover@chbe.gatech.edu; NVH: hud@chemistry.gatech.edu

DGL: Voice (+1)404 727-9348; Fax (+1)404 727-8658; E-mail dlynn2@emory.edu

\*\* We dedicate this manuscript to Albert Eschenmoser and his foundational contributions on which much of this work is based, thank Dr. Ram Krishnamurthy of the Scripps Institute for valuable insight and acknowledge the NIH RR12723, Seaver Foundation, a Georgia Institute of Technology/Emory University FAME Center Initiative, and the NSF/NASA Center for Chemical Evolution for financial support.

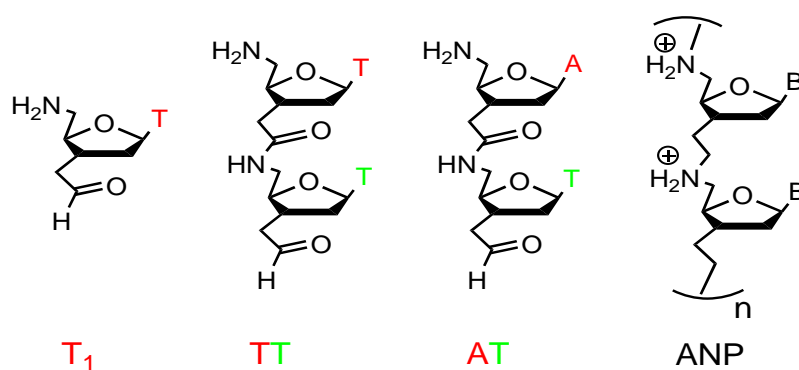
**Abstract** - A self-organizing process is described that combines DNA template association thermodynamics and kinetic reductive amination to translate polymer sequence information into amine nucleoside polymers. The developed kinetic model analyses allowed this process to be extended to the translation of templates as long as catalytic ribozymes.

## INTRODUCTION

The Watson-Crick structure of DNA reported in 1953 elegantly revealed the central role of template-directed polymerization in Darwinian evolution.<sup>1</sup> Within that same month, publication of the Urey-Miller experiment outlined a strategy for defining a chemical inventory for the emergence of life.<sup>2</sup> Building on

these foundations, Albert Eschenmoser and his colleagues<sup>3-6</sup> have developed syntheses of potentially natural alternative structures and provided deep insight into a chemical etiology of the nucleic acid duplex. We now understand the natural bases to be versatile molecular recognition elements capable of many self- and nonself-pairing motifs<sup>7</sup> with the sugar phosphate backbone contributing significantly to natural pairing arrangements. However, phosphate-modified oligonucleotides also form stable duplexes.<sup>8-10</sup> And sugars other than ribose, some with much simpler carbon skeletons,<sup>5,6</sup> also stabilize Watson-Crick base pairs. Even backbones containing no phosphate and no sugar form stable DNA/RNA heteroduplexes.<sup>11</sup> Eschenmoser has argued that the native backbone is not a unique solution, in fact evolutionary pressures have not even maximized duplex stability, rather there exists a broad region of structural space accessible to information encoding templates. One must wonder whether this argument then implies that template-directed replicative polymerization can also be readily extended into molecular skeletons distinct from the present day nucleic acids.<sup>12</sup>

In their simplest form, natural polymerases employ a feedback circuit, coupling the thermodynamic condensation of monomers on the template with a subsequent kinetic trap that hydrolyzes the pyrophosphate product. This process has been simplified by combining reversible imine condensation on a template with kinetic reductive amination for controlling polymer growth.<sup>13-16</sup> Specifically, thymidine carrying a 5'-amine and a 3'-acetaldehyde, 5'-H<sub>2</sub>N-dT-3'-CH<sub>2</sub>CHO, T<sub>1</sub>, provides the polymerization monomer (Figure 1). While T<sub>1</sub> can form intramolecular imines, no reaction is detected under mild reducing conditions in aqueous environments. In contrast, T<sub>1</sub> was reductively oligomerized on a (dAp)<sub>8</sub> DNA template to give a single amine octamer product 5'-H<sub>3</sub>N<sup>+</sup>-(dT-3'-CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub><sup>+</sup>)<sub>7</sub>-dT-3'-CH<sub>2</sub>CHO, T<sub>8</sub> (Table 1).<sup>14</sup>



**Figure 1.** Monomer T<sub>1</sub>, amide dimers TT and AT, and the reductive amination product amine nucleoside polymer (ANP).

As expected from the modified sugar backbone, ligation experiments established antiparallel substrate/template association and robust sequence specific reactions on short DNA templates.<sup>14</sup> However, the progressive accumulation of amine nucleoside polymer (ANP) product intermediates over time, T<sub>2</sub>

and T<sub>4</sub>, with no trimer, pentamer, hexamer, or heptamer products, was striking. We have wondered whether these results suggest a general strategy to produce novel gene-length polymers through sequence and chain-length specific step growth template-directed polymerization.

**Table 1.** Yield of reaction products with time

Time of Analysis*	Monomer (T <sub>1</sub> , %)	Dimer (T <sub>2</sub> , %)	Tetramer (T <sub>4</sub> , %)	Octamer (T <sub>8</sub> , %)
Time zero	100	0	0	0
6 h	0	99	0	0
24 h	0	15	80	0
48 h	0	0	22	70
60 h	0	0	0	91

\* HPLC quantification of template-directed reductive polymerization of (T)<sub>1</sub> on a (dA<sub>p</sub>)<sub>8</sub> template with 20 M eq. of NaBH<sub>3</sub>CN. Additional reductant is added at 48 h to complete the reaction.

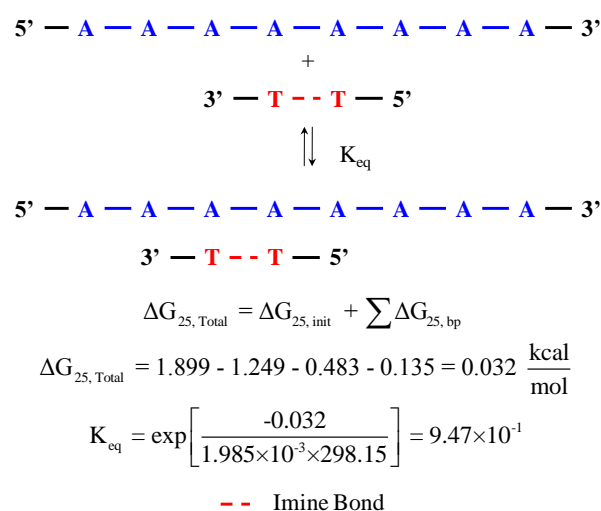
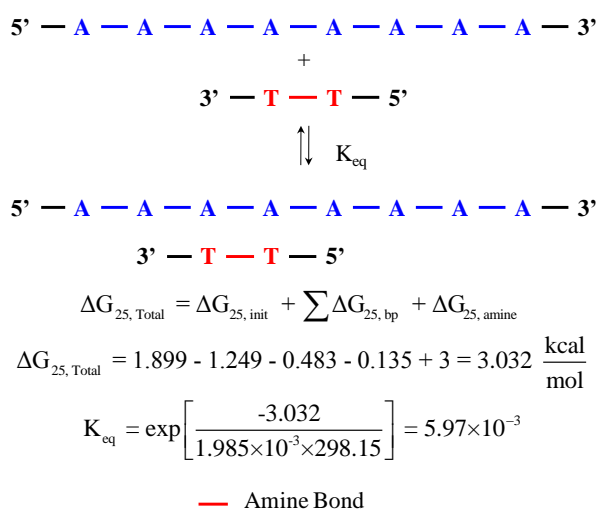
## RESULTS

*Modeling template-directed polymerization reactions.* Step-growth reaction dynamics are clearly controlled by the polymeric template. With an *n*-unit template that forms preferential anti-parallel complexes with complementary monomers, there are *n*-1 unique imine dimer-binding sites, *n*-3 imine tetramer sites, and *n*-7 octomer sites. Such a progressive reduction in the number of sites will impact the imine concentration and the reduction rates. Maybe more importantly, the amine backbone oligomers are known to associate more weakly to a DNA template than the imine,<sup>13,16</sup> further impacting the concentration of template-bound product early in oligomerization. Accordingly, we have constructed a model for the reaction using nearest-neighbor thermodynamic parameters for DNA Watson-Crick base pairs<sup>17</sup> to estimate equilibrium constants, and Table 2 lists the parameters used in the calculations.

We have modeled the association of the rigid imine oligomer as equivalent to a DNA oligomer with the same nucleotide sequence.<sup>18</sup> To emulate the weaker association when an amine bond is present,<sup>18</sup> an additional Gibbs energy penalty,  $\Delta G_{25,\text{amine}}$ , is included for each amine bond as illustrated in Figure 2. The equilibrium constant for an oligomer with amine bonds is estimated using the thermodynamic parameters in Table 2, along with the additional  $\Delta G_{25,\text{amine}}$  parameter for each amine bond present in the oligomer backbone. Equilibrium constants are listed in Table 3 for various species with several selected values of  $\Delta G_{25,\text{amine}}$ .

**Table 2.** Nearest-neighbor thermodynamic parameters for DNA Watson-Crick pairs<sup>17</sup> at a temperature of 25 °C

Sequences		$\Delta G_{25}$ (kcal mol <sup>-1</sup> )
Initiation Sequence		1.899
Propagation Sequences	A A T T	-1.249
	A T T A	-1.118
5' Dangling End Sequences	A A T	-0.483
	T A T	-0.949
3' Dangling End Sequences	A A T	-0.135
	A T T	-0.022
	T A A	-0.488



**Figure 2.** Equilibrium constants for templated-bound products of modified thymidine monomer T<sub>1</sub> on a (dAp)<sub>8</sub> DNA template by reductive amination using  $\Delta G_{25, \text{amine}} = 3.0$  kcal mol<sup>-1</sup>. The (dAp)<sub>8</sub> DNA sequences are in blue, and the polymer products in red. T<sub>2</sub> oligomers are connected by amine or imine backbones. The equilibrium constants reflect the weaker association of the amine backbone on the template because of the proposed thermodynamic amine parameter.

When  $\Delta G_{25, \text{amine}}$  is equal to or greater than 3.0 kcal mol<sup>-1</sup>, the dimer amine (T<sub>2</sub>-a) is 20-fold less likely to be on the template compared to the T<sub>1</sub> monomer. Under those conditions, virtually all T<sub>1</sub> monomers would react to form (T<sub>2</sub>-a) oligomers more quickly than a (T<sub>2</sub>-a) oligomer could react with a third T<sub>1</sub> monomer to form a trimer amine-imine oligomer (T<sub>3</sub>-ai) on the template. Once only dimers are present

in the system, pentamers and heptamers cannot form. However, from ( $T_2$ -a) dimers, tetramers are eventually formed on the template and, like the dimer-monomer relationship, tetramers ( $T_4$ -aaa) will desorb rapidly from the template before they can react with an additional ( $T_2$ -a) dimer to form a hexamer. The quantitative modeling summarized in Table 3 supports the hypothesis that the weak interaction of the amine backbone oligomer with the DNA template prevents trimer, pentamer, hexamer and heptamer from forming via template-directed polymerization.

**Table 3.** Equilibrium constants ( $K_{eq}$ ) for different template-bound products on a (dAp)<sub>8</sub> DNA template using the proposed thermodynamic amine parameter,  $\Delta G_{25,amine}$  at 25 °C

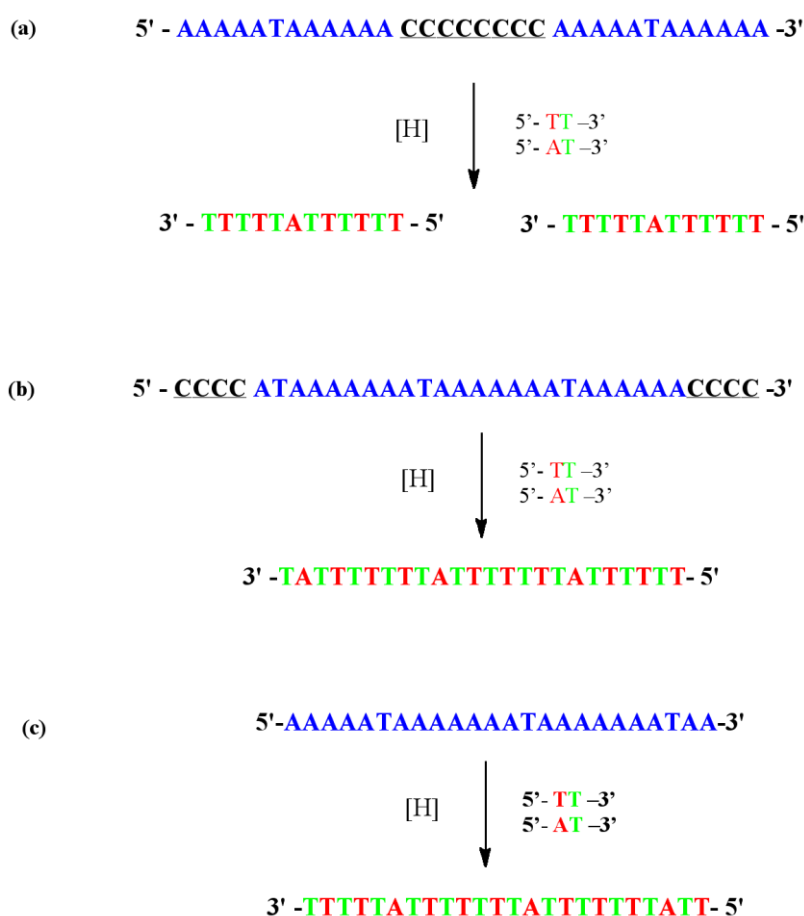
Oligomer	Backbone Structure*	$\Delta G_{25,amine}$ (kcal mol <sup>-1</sup> )			
		1.0	2.0	3.0	4.0
T <sub>1</sub>	---	$1.15 \times 10^{-1}$	$1.15 \times 10^{-1}$	$1.15 \times 10^{-1}$	$1.15 \times 10^{-1}$
T <sub>2</sub>	i	$9.47 \times 10^{-1}$	$9.47 \times 10^{-1}$	$9.47 \times 10^{-1}$	$9.47 \times 10^{-1}$
T <sub>2</sub>	a	$1.75 \times 10^{-1}$	$3.23 \times 10^{-2}$	$5.97 \times 10^{-3}$	$1.10 \times 10^{-3}$
T <sub>3</sub>	a - i	$1.44 \times 10^0$	$2.66 \times 10^{-1}$	$4.92 \times 10^{-2}$	$9.09 \times 10^{-3}$
T <sub>3</sub>	a - a	$2.66 \times 10^{-1}$	$9.09 \times 10^{-3}$	$3.10 \times 10^{-4}$	$1.06 \times 10^{-5}$
T <sub>4</sub>	a - a - a	$4.06 \times 10^{-1}$	$2.56 \times 10^{-3}$	$1.61 \times 10^{-5}$	$1.02 \times 10^{-7}$

\*a: Amine bond, i: Imine bond

The experimental work reported by Zhan *et al.*<sup>18</sup> showed that adding a single amine bond can create a 10<sup>6</sup>-fold difference in  $K_{eq}$ , resulting in the value of  $\Delta G_{25,amine} = 8.75$  kcal mol<sup>-1</sup>. Such significant energy differences suggest that great control can be exerted over the reaction and that even longer sequences, approaching the complexity of simple ribozymes, might be used as catalytic templates.

*Translating longer sequences.* Recognizing that homopurine and homopyrimidine sequences could form stable triple helix structures,<sup>11</sup> blocks of non-complementary cytosine bases were inserted within and around the reading frame (Figure 3). The two 12-mer templates in Figure 3a were translated with 5'-amino-3'-acetaldehyde amide dimers TT and AT, and while the 12-mer amide/amine nucleoside polymer (ANP) was the largest molecular weight product detected, several truncation products were apparent. With cytosine tracks capping both 5'- and 3'- ends (Figure 3b) however, the 24-mer ANP product was produced cleanly. Moreover, this product was identical in composition and rate of formation to that obtained with the simpler 24-mer templates shown in Figure 3c.

Previous results had established that the homothymidine ANP amide octamer prepared by solid phase synthesis forms only 1:1 duplexes with (dAp)<sub>8</sub>, and the complex melts close to room temperature in the absence of added salt.<sup>19</sup> The octamer at this temperature may then represent a transition point for polymerization kinetics on the template, and the difference in the effectiveness of the 12-mer vs. 24-mer templates could be limited by complete eight base concatamers on the template. To test this idea, the 32-mer DNA template strand, 5'-dAAAAATAAAAAATAAAAAATAAAAAATAA-3' (in blue), was designed to include four identical octamer blocks. As shown in Figure 4, polymerization of the amide dimers TT and AT on this template gave the 32-mer-ANP as the sole product with no detectable intermediates.



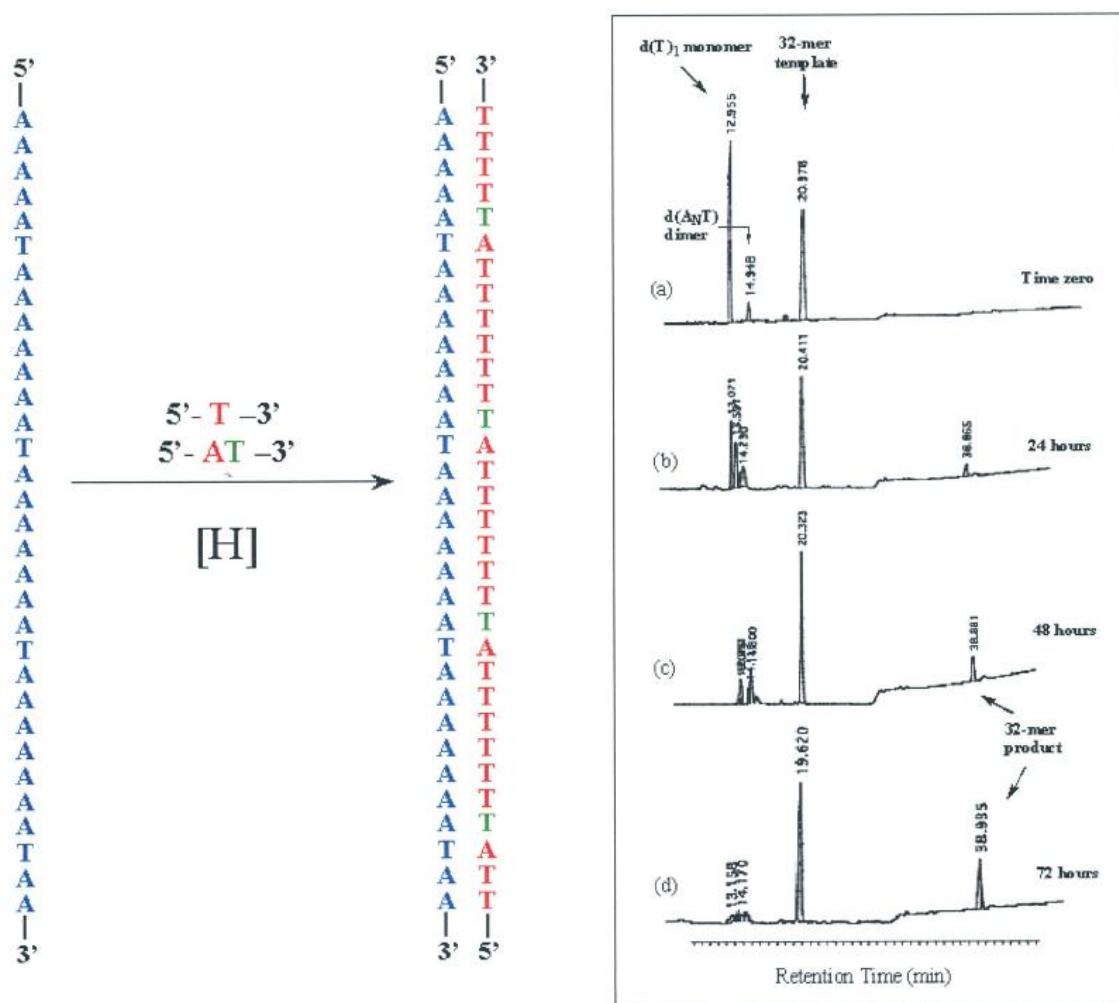
**Figure 3.** Translating longer DNA templates (a) Templates contain a central oligo C block; (b) terminal C blocks; and (c) a DNA 24-mer template. The ApT DNA sequences are highlighted in blue, the non-complementary regions are in black and underlined, and the product bases in green represent a 5' amide. [H] designates reductive amination via NaBH<sub>3</sub>CN. MALDI-TOF (THAP/citrate matrix): 12-mer product in (a): m/z 3124.47 (M+2H)<sup>+</sup>, calc. C<sub>144</sub>H<sub>193</sub>N<sub>39</sub>O<sub>41</sub>, 3124.46. 24-mer product in (b, c): m/z 6284.03 (M+2Na)<sup>+</sup>, calc. C<sub>288</sub>H<sub>381</sub>N<sub>81</sub>O<sub>79</sub>Na<sub>2</sub> 6283.81.

Most importantly, no truncation products were detected within the resolution of HPLC and MALDI-TOF mass spectrometry analyses. While sequencing methods of these polymers are still being developed, a



Our thermodynamic modeling supports this conclusion. At 25 °C, we compute a  $K_{eq} = 1.38$  for the AT amide dimer on the template, using the parameters in Table 2. More surprisingly,  $K_{eq} = 2.08$  for the TT amide dimer when it is located on the template such that the T base on the DNA template provides a stabilizing 3' dangling end interaction. Both values are higher than the nominal value of  $K_{eq} = 0.947$  seen in Table 3 for the TT dimer on the octameric (dAp)<sub>8</sub> template.

Our modeling also provides an explanation for the faster polymerization in Figure 4 relative to Figure 5, as is shown in Table 4. Because the TT dimer (with amide bond) should associate more strongly with the template than the T<sub>1</sub> monomer, it is expected that polymerization would proceed more rapidly using the TT dimer. Indeed, employing T<sub>1</sub> and AT substrates requires longer reaction time for the polymerization, but also gives the single 32-mer ANP product with no detectable intermediates. Most critically, the backbones generated by the two reactions are different, but the ANP base sequences are identical.



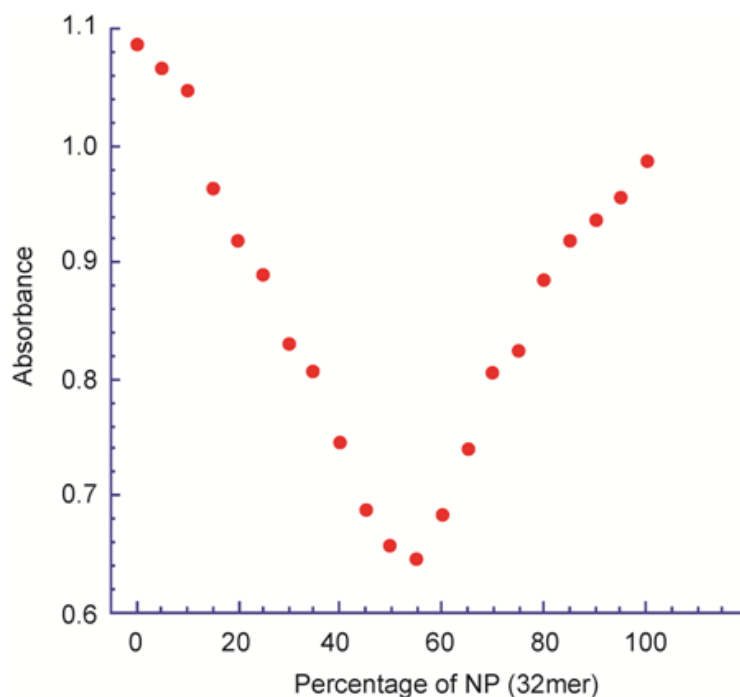
**Figure 5.** HPLC analysis of the 32-DNA catalyzed by 5'-T-3' and 5'-AT-3' polymerization. (a) time zero; (b) 24 h; (c) 48 h; (d) 72 h. MALDI-TOF (THAP/citrate matrix): 32-mer product  $m/z$  8142.34 ( $M+2H$ )<sup>+</sup>, calc. C<sub>384</sub>H<sub>534</sub>N<sub>108</sub>O<sub>93</sub>, 8142.04.

**Table 4.** Yield of 32-mer NP product with time

Substrates*	0 time	12 h	24 h	48 h	72 h
TT, AT	0 %	5%	18%	81%	nd
T <sub>1</sub> , AT	0%	nd	6%	27%	76%

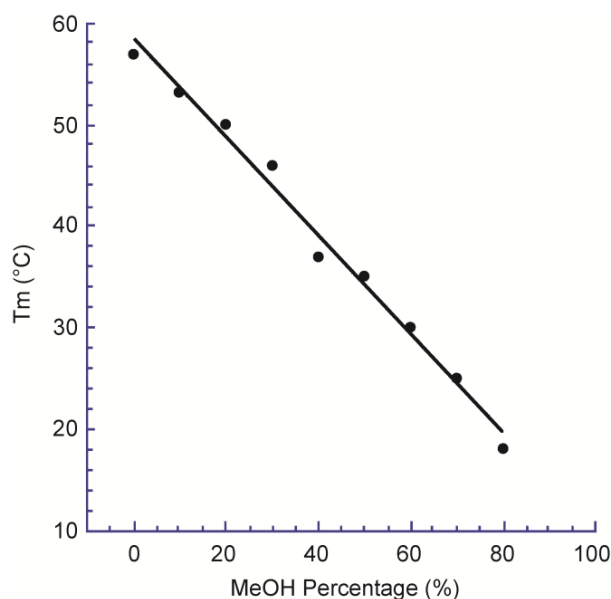
\*HPLC quantification of NP from the template-directed reductive polymerization of the indicated substrates on the 32-mer DNA template in 20 M eq. of NaBH<sub>3</sub>CN. Additional reductant was added at 48 h. nd, not determined.

*The ANP/DNA molecular complex.* To better understand the physical properties of the ANP oligomers, the 32-mer DNA/DNA duplex with the same sequence had a measured T<sub>m</sub> of 44 °C in 500 mM NaCl. Under similar conditions but in the absence of salt, the ANP/DNA 1:1 complex prepared with T<sub>1</sub> and AT melted cooperatively at 75 °C, and the complex prepared with TT and AT substrates melted at > 87 °C. Consistent with the shorter length sequences, the more rigid amide backbones show increased thermal stability. The hypochromic shift of the base electronic transitions used for the melting analysis, typically around 40% in the DNA double helix, is almost 70% in the ANP/DNA complexes, suggesting similar base stacking associations. Indeed, Job plot analyses of various mole fraction ratios of the product 32-mer ANP and DNA template strands show a hypochromic minimum of a 1:1 complex (Figure 6).



**Figure 6.** Absolute absorbance plotted as a function of different mole fraction of the DNA template with the 32-mer ANP product prepared in Figure 2. Conditions: 1 μM for total concentration, 10 mM NaH<sub>2</sub>PO<sub>4</sub> buffer (pH=6), and measured at 260 nm.

Consistent with the previously observed insensitivity to salt, the ANP/DNA complex remains stable in low dielectric media. Figure 7 shows that the melting temperature of the 1:1 32-mer TT/AT ANP/DNA complex is linearly dependent on added MeOH, at least up to 80% where the phosphate buffer precipitates. Each melting event in this series appears to follow a simple two-state, strand/complex equilibrium.



**Figure 7.** Dependence of ANP/DNA duplex melting temperature ( $T_m$ ) as a function of added MeOH.

## DISCUSSION

While additional experiments will be required to build structural models for the ANP/DNA complex, these observations are most consistent with a duplex structure similar to that seen for native nucleic acids. Such stable, organic solvent soluble nucleic acid hybrid materials are nevertheless physically distinct from present day biopolymers, and unique biotechnological applications ranging from gene delivery, control of gene expression, and highly specific diagnostic applications may now be possible with these materials.

More generally, it is now clear that reductive amination polymerization templated by DNA is generally effective,<sup>20,21</sup> and can indeed template the sequence specific polymerization of macromolecules approximating the length of the smallest catalytic ribozymes.<sup>22</sup> This reaction is however more analogous to DNA translation to a different backbone, and greater understanding of the thermodynamics of template association and reductive amination kinetics will be required to extend to longer DNA reads. Even at this point however, several advantages are apparent. For example, the chain-growth processes catalyzed by the natural polymerases on homopolymer templates can result in catalyst slippage and lead

to significant error rates. In contrast, the step-growth reaction reported here creates a self-organizing process that allows homopolymers to be translated efficiently.

And finally, alternate reactions and strategies for molecular information storage and transfer sets the stage for a synthetic chemical evolution. The demonstration that these amine backbone structures are capable of catalyzing phosphodiester ligation,<sup>23</sup> suggests that the ANPs could also serve as transitional catalysts for even more complex evolutionary processes. In short, these methods take us one step closer to providing a robust and mild chemical method for reading the sequence information of DNA into alternate backbones, building on and extending the remarkable reactions that must have enabled the origins of chemical evolution.

## REFERENCES AND NOTES

*Polymerization Reactions.* The substrate and the template were mixed in a stoichiometric ratio and diluted by ddH<sub>2</sub>O to 20  $\mu$ L. The final concentration of the reaction mixtures was 0.5 mM for the template, and the substrate total concentration is always a complement to the template in a molar 1:1 ratio. The reaction mixture prepared was heated to 75 °C for 2-3 min., then cooled to 4 °C for 3 h. NaBH<sub>3</sub>CN (20 eq.) was added at room temperature and the reaction mixture was vigorously vortexed throughout the reaction. Extra portions of NaBH<sub>3</sub>CN were added towards the end of some reactions to ensure the completion. The mixture was sampled (0.5  $\mu$ L) at the indicated time points for HPLC analysis: Rainin HPXL, Phenomenex Prodigy 5 analytical ODS(2) C18; Rainin Dynamax UV detector at 260 nm, and eluted with MeOH in H<sub>2</sub>O: 0% -5% from 0-8 min, 5% -20% from 8 -9 min, 20% -35% from 9 -24 min, 35% -60% from 24 -25 min, 60% -100% from 25 -50 min.

*Job plots.* Two equal molar aqueous stock solutions at pH 6 in 10 mM phosphate buffer, one of the DNA template and the other of the 32-mer NP, were mixed in the indicated ratios so that the sample solution was maintained at 500  $\mu$ L. Each of the samples was heated to 75 °C for 3-5 min and incubated at 4 °C for 3 h prior to analysis at 260 nm. The absorbance of the blank was subtracted from the total absorbance of each sample.

*Thermal melting.* Stock solutions were prepared at 1  $\mu$ M of total oligonucleotide concentration, buffered with 10 mM NaH<sub>2</sub>PO<sub>4</sub> to pH 6, and mixed to a total volume 250  $\mu$ L. The native duplex sample required 500 mM NaCl to give adequate melting analyses. Each sample was sealed in the UV cuvette with Teflon tape to prevent the solvent evaporation at high temperatures. The peltier thermal controller was ramped from 5 °C to 95 °C and the UV absorbance was monitored at 260 nm. For MeOH denatured melting

experiments, each sample was prepared as above with the indicated MeOH percentage.

*Cited References.*

1. J. D. Watson and F. H. C. Crick, [\*Nature\*, 1953, \*\*171\*\*, 737.](#)
2. S. L. Miller, [\*Science\*, 1953, \*\*117\*\*, 528.](#)
3. A. Eschenmoser, [\*Science\*, 1999, \*\*284\*\*, 2118.](#)
4. A. Eschenmoser, [\*Tetrahedron\*, 2007, \*\*63\*\*, 12821.](#)
5. K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy, and A. Eschenmoser, [\*Science\*, 2000, \*\*290\*\*, 1347.](#)
6. A. Eschenmoser, [\*Origins of Life Evol. B.\*, 2004, \*\*34\*\*, 277.](#)
7. P. Hobza and J. Šponer, [\*Chem. Rev.\*, 1999, \*\*99\*\*, 3247.](#)
8. A. De Mesmaeker, R. Haener, P. Martin, and H. E. Moser, [\*Acc. Chem. Res.\*, 1995, \*\*28\*\*, 366](#) and the references therein.
9. A. De Mesmaeker, K-H. Altmann, A. Waldner, and S. Wendeborn, [\*Curr. Opin. Struct. Biol.\*, 1995, \*\*5\*\*, 343.](#)
10. S. A. Benner and D. Hunter, [\*Bioorganic Chem.\*, 2002, \*\*30\*\*, 62.](#)
11. P. E. Nielsen, M. Egholm, R. H. Berg, and O. Buchardt, [\*Science\*, 1991, \*\*254\*\*, 1497.](#)
12. G. F. Joyce, [\*Nature\*, 2002, \*\*418\*\*, 214.](#)
13. Y. Gat, J. Ye and D. G. Lynn, Template-Directed Ligation: Towards the synthesis of sequence specific polymers, in *Templated Organic Synthesis*, Ed. By P. J. Stang and F. Diederich, 2000, pp. 133-58.
14. X. Li, Z.-Y. J. Zhan, R. Knipe, and D. G. Lynn, [\*J. Am. Chem. Soc.\*, 2002, \*\*124\*\*, 746.](#)
15. X. Li and D. G. Lynn, [\*Angew. Chem.\*, 2002, \*\*114\*\*, 4749.](#)
16. Z.-Y. J. Zhan and D. G. Lynn, [\*J. Am. Chem. Soc.\*, 1997, \*\*119\*\*, 12420.](#)
17. J. SantaLucia, Jr. and D. Hicks, [\*Annu. Rev. Biophys. Biomol. Struct.\*, 2004, \*\*33\*\*, 415.](#)
18. Z.-Y. J. Zhan, J. Ye, X. Li, and D. G. Lynn, [\*Current Org. Chem.\*, 2001, \*\*5\*\*, 885.](#)
19. J. C. Leitzel and D. G. Lynn, [\*Chem. Rec.\*, 2001, \*\*1\*\*, 53.](#)
20. D. M. Rosenbaum and D. R. Liu, [\*J. Am. Chem. Soc.\*, 2003, \*\*125\*\*, 13924.](#)
21. Y. Brudno, M. E. Birnbaum, R. E. Kleiner, and D. R. Liu, [\*Nature Chemical Biology\*, 2010, \*\*6\*\*, 148.](#)
22. J. A. Doudna and T. R. Cech, [\*Nature\*, 2002, \*\*418\*\*, 222.](#)
23. J. Ye, Y. Gat, and D. G. Lynn, [\*Angew. Chem. Intl. Ed.\*, 2000, \*\*39\*\*, 3641.](#)