Supporting Information

Efficient and Sequence-Specific DNA-Templated Polymerization of PNA Aldehydes

Daniel M. Rosenbaum and David R. Liu*
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Oligonucleotide synthesis, purification and sequences. DNA oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8090 DNA synthesizer using standard phosphoramidite protocols. All reagents were purchased from Glen Research. The templates for all polymerization experiments were synthesized with a 5'-MMT-amino-dT phosphoramidite as the 5' terminal monomer coupled to the oligo. For the oligonucleotides used in Figures 1 and 2, purification of DNA templates was carried out with the following steps: 1) Exocyclic amine deprotection with 1:1 ammonium hydroxide:methylamine; 2) reverse-phase HPLC purification of the MMT-on fraction; 3)MMT cleavage; 4) HPLC purification of the MMT-deprotected product. The oligonucleotides used in Figure 3 were deprotected with 1:1 ammonium hydroxide:methylamine, and then subjected to OPC cartridge purification (Applied Biosystems).

The sequences of the templates are listed below, where N = 5'-amino-dT:

Figure 1
Lane 1: NCGAATTCGTACGAATTCGAAGTCAGTCAGGCAGTCAGTCTTTT
Lane 2: NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCTTTT
Lane 3: NCGAATTCGTACGAATTCGAATGCAGTCAGTCAGTCAGTCTTTT
Lane 4: NCGAATTCGTACGAATTCGAAGTCATGCAGTCAGTCAGTCTTTT
Lane 5: NCGAATTCGTACGAATTCGAAGTCAGTCATTCAGTCAGTCTTTT
Lane 6: NCGAATTCGTACGAATTCGAAGTCAGTCATGCAGTCAGTCTTTT
Lane 7: NCGAATTCGTACGAATTCGAAGTCAGTCATCCAGTCAGTCTTTT
Lane 8: NCGAATTCGTACGAATTCGAAGTCAGTCAGGCAGTCAGTCTTTT
Lane 9: NCGAATTCGTACGAATTCGAAGTCAGTCAGCCAGTCAGTCTTTT
Lane 10: NCGAATTCGTACGAATTCGAAGTCAGTCACTCAGTCAGTCTTTT
Lane 11: NCGAATTCGTACGAATTCGAAGTCAGTCACGCAGTCAGTCTTTT
Lane 12: NCGAATTCGTACGAATTCGAAGTCAGTCACCCAGTCAGTCTTTT
Lane 13: NCGAATTCGTACGAATTCGAAGTCAGTCAGTCATGCAGTCTTTT
Lane 14: NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCATGCTTTT
Lane 15: NCGAATTCGTACGAATTCGAAATGAATGAATGAATGAATGTTTT

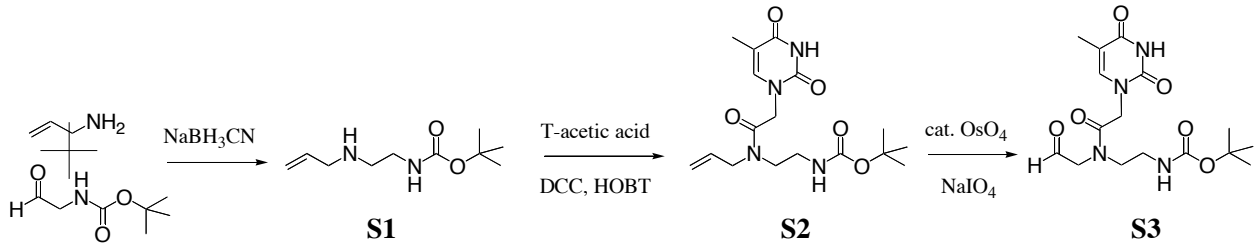
Figure 2
Lanes 1-2:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 3-4:
NCGAATTCGTACGAATTCGAATGCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 5-6:
NCGAATTCGTACGAATTCGAAGTCAGTCATGCAGTCAGTCAGTCAGTCAGTCAGTCAGTC
TTTT

Lanes 7-8:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCATGCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 9-10:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCAGTCATGCAGTCAGTCAGTC
TTTT
Lanes 11-12:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCATGCAGTC
TTTT

Figure 3, top gel
Lanes 1-2:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 3-6:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCATTCAGTCAGTCAGTCAGTCAGTCT
TTT
Lanes 7-10:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCACTCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 11-14:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCATCCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 15-18:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCACCCAGTCAGTCAGTCAGTCAGTC
TTTT

Figure 3, bottom gel
Lanes 1-2:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 3-6:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGCCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 7-10:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCATGCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 11-14:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCACGCAGTCAGTCAGTCAGTCAGTC
TTTT
Lanes 15-18:
NCGAATTCGTACGAATTCGAAGTCAGTCAGTCAGTCAGGCAGTCAGTCAGTCAGTCAGTC
TTTT

Synthesis of Boc-t-CHO PNA monomer aldehyde (S3)



The basic protocol for formation of the allyl-substituted secondary amine (**S1**) and its coupling to thymine-1-acetic acid was adapted from Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P.E. *Tetrahedron Letters* **1998**, 39, 4707.

Preparation of allyl-substituted secondary amine (**S1**)
All chemicals, unless otherwise noted, were purchased from Aldrich. Allylamine (0.493 ml, 6.3 mmol) and Boc-aminoacetaldehyde (1.05 g, 6.3 mmol) were dissolved together in 40 ml dry methanol. After 30 minutes at room temperature, 410 mg $NaBH_3CN$ (6.5 mmol) and 0.375 ml acetic acid were added to the solution. The reaction was allowed to proceed 3 h at room temperature, and the solvent was evaporated. The residue was resuspended in 50 ml saturated sodium bicarbonate, and the product was extracted with 3 x 50 ml ethyl acetate, washed with 2x 20 ml brine, and purified by flash chromatography with a solvent system of 3:1 ethyl acetate:methanol (R_f = 0.3). Yield = 25% (320 mg). ^{1}H NMR (500 MHz, CD_3OD) □ 1.5 (s, 9H), 3.1 (m, 2H), 3.5 (m, 2H), 3.6 (m, 2H), 5.25 (m, 1H), 5.45 (m, 1H), 5.85 (m, 1H).

Preparation of allyl-substituted thymine PNA monomer (**S2**)
Thymine-1-acetic acid(1.56 g, 8.5 mmol), DCC (2.0 g, 9.5 mmol), and HOBT (1.5 g, 8.5 mmol) were dissolved in 30 ml dry DMF at room temperature After 5 minutes, upon observation of DCU precipitation, allyl-substituted secondary amine (**S1**) (1.7 g, 8.5 mmol) and 1.2 ml triethylamine were added to the solution. After 24 hrs. at room temperature, DMF was evaporated, residue was redissolved in 50 ml ethyl acetate, washed with 2 x 25 ml saturated sodium bicarbonate, washed with 2 x 25 ml brine, and purified by flash chromatography with ethyl acetate (R_f = 0.5). Yield = 60% (1.4 g). ^{1}H NMR (500 MHz, CD_3OD) □ 1.4 (s, 9H), 1.9 (s, 3H), 3.25 (m, 2H), 3.45 (m, 2H), 4.0 (m, 2H), 4.5 (m, 2H), 5.2 (m, 1H), 5.3 (m, 1H), 5.8 (m, 1H), 7.0 (m, 1H). APCI MS: Expected (MH^+)- 367.2. Observed- 366.9.

The protocol for OsO_4-catalyzed cleavage of the allyl group to generate the aldehyde **S3** was adapted from Luo, P.; Leitzel, J.C.; Zhan, Z.-Y. J.; Lynn, D.G. *J. Am. Chem. Soc*. **1998**, 120, 3019.

Preparation of Boc-t-CHO PNA monomer aldehyde (**S3**)
Allyl-substituted thymine PNA monomer (**S2**)(1.4 g, 3.8 mmol) was dissolved in 60 ml 2:1 dioxane:H_2O. 140 □l 4% aqueous OsO_4 was added to the solution at room temperature and a light brown color appeared after ~15 minutes. Over the next hour, 1.2 g solid $NaIO_4$ was added in small portions to the solution. A precipitate was formed, and the reaction was allowed to proceed for 24 hours at room temperature The precipitate was then filtered off, and the product was extracted with 3 x 60 ml ethyl acetate, washed with 2 x 50 ml brine, and purified by prep TLC with a solvent system of 9:1 ethyl acetate:methanol (R_f = 0.5). Yield = 34% (0.475 g). (1.4 g). ^{1}H NMR (500 MHz, CD_3OD) □ 1.4 (s, 9H), 1.9 (s, 3H), 3.4 (m, 2H), 3.7 (m, 2H), 3.8 (m, 2H), 4.55 (m, 2H), 6.95 (m, 1H), 8.9 (m, 1H). ES MS: Expected (MH^-)- 367.2. Observed- 367.1.

Preparation of T-aldehyde resin. 1g aminomethyl styrene tentagel resin (loading = 0.29 mmol/g) was swelled in 10 mL DMF for 10 minutes, and the solvent removed. FMOC-threonine was activated as follows: 0.495 g FMOC-Thr-OH (1.45 mmol), 0.552 g HATU (1.45 mmol), 0.492 mL DIPEA (2.90 mmol), and 0.164 mL lutidine (1.45 mmol) were dissolved in 10 mL DMF, stirred for 7 min at room temperature, and then added to the resin. The resin plus activated FMOC-threonine was agitated in a peptide synthesis vessel with nitrogen bubbling for 3 h at room temperature, washed with DMF, washed with DCM, washed with methanol, and desiccated. The resin was then re-swelled in DMF, and the solvent removed. The FMOC group was removed by resuspending resin in 10 mL 20% piperidine in DMF and agitating 10 min. at room temperature (repeated twice). The FMOC deprotection eluates were collected and threonine loading was quantitated by measuring A_{300} of a small fraction, indicating near quantitative substitution of the resin (0.29 mmol threonine per g resin). The resin was washed with DMF, DCM, and MeOH, and then resuspended in a solution of 0.34 g Boc-t-CHO (**S3**) plus 0.1 mL DIPEA in 10 mL methanol. The suspension was stirred at 60 °C for 3 h, washed with a large volume of methanol, and desiccated. It was stored at 4 °C, and was stable for up to 6 months.

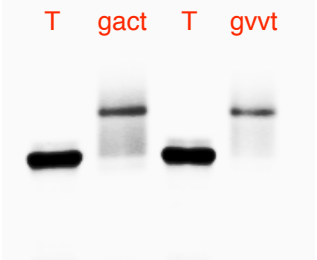
Peptide synthesis using the T-aldehyde resin. Peptide aldehyde synthesis was carried out essentially as in Ede and Bray, *Tetrahedron Letters* **1997**, 38, 7119. FMOC-protected PNA monomers (with Bhoc-protected exocyclic amines) were purchased from Applied Biosystems. A representative PNA peptide aldehyde synthesis is described below, using the NH_2-gact-CHO sequence. 100 mg T-aldehyde resin was swelled with 5 mL DMF for 10 min, and washed with DCM. The Boc group was deprotected by adding 5 mL [TFA + 5% m-cresol], and agitating resin with nitrogen bubbling for 10 min (repeated twice). Resin was washed with DCM, DMF, and pyridine. The FMOC-c-OH PNA monomer was activated as follows: 80.7 mg FMOC-c-OH, 44 mg HATU, 40 μL DIPEA and 13.5 μL were dissolved in 2.5 mL DMF, and allowed to react for 7 min at room temperature, after which a yellow color was observed. This mixture was then added directly to the Boc-deprotected T-aldehyde resin in a peptide synthesis vessel, and the suspension was agitated with nitrogen bubbling for 30 min at room temperature The resin was then drained and washed with a large volume of DMF. FMOC deprotection was carried out by adding 5 mL [20% piperidine in DMF], and agitating resin for 5 min at room temperature (repeated twice), followed by washing with a large volume of DMF. The PNA monomer coupling and FMOC deprotection cycle was then carried out with FMOC-a-OH (Activation: 83.5 mg FMOC-a-OH + 44 mg HATU + 40 μL DIPEA + 13.5 μL lutidine), followed by FMOC-g-OH (Activation: 85.3 mg FMOC-g-OH + 44 mg HATU + 40 μL DIPEA + 13.5 μL lutidine). After the final FMOC deprotection, the exocyclic amines were deprotected by resuspending the resin in 5 mL [TFA + 5% m-cresol], and agitating 10 min at room temperature (repeated twice), after which the resin was washed with DCM and methanol. Cleavage from the resin was carried out by resuspending the resin in 5 mL 60:40:1 CH_3CN:H_2O:TFA, and stirring at 60 °C for 1 h. The cleavage cocktail was filtered away from the resin, frozen, and lyophilized overnight. The resulting residue was resuspended in 0.1% TFA in H_2O and subjected to semipreparative reverse-phase HPLC (4 mL/min) with a gradient of [0.1% TFA in H_2O] → 70:30 [0.1% TFA in H_2O]: CH_3CN over 30 minutes. The primary peak fraction eluted at 14-15 min; it was lyophilized and resuspended in pure H_2O. The concentration was determined by measuring A_{260} and assuming an extinction coefficient of 38,800 $L \bullet mol^{-1} \bullet cm^{-1}$, as in the analogous DNA sequence. A small aliquot of this sample was subjected to positive ion mode electrospray mass spectrometry. The expected mass of the product PNA aldehyde was 1086 D. The observed mass of the collected fraction was 1086.6. All other PNA peptide aldehydes were synthesized in an analogous fashion to NH_2-gact-CHO. The expected and observed masses from the positive ion electrospray mass spectra are summarized in the following table:

Sequence	Expected mass	Observed mass*
gaat	1109	1109.4
gagt	1125	1125.4
ggat	1125	1125.4
gggt	1141	1141.4
ggct	1101	1101.4
gcat	1085	1085.4
gcgt	1101	1101.4
gcct	1061	1061.4

*Peaks corresponding to the hydrate of the C-terminal aldehyde were also observed in the mass spectrum of each purified peptide.

DNA-templated reductive amination polymerization protocol. In a typical DNA-templated PNA aldehyde polymerization experiment, 20 pmol 5'-amino-terminated DNA template and the relevant PNA peptide aldehydes (400 pmol total for experiments in Figure 1; 800 pmol total for experiments in Figures 2 and 3) were mixed together with 25 μL 2 M NaCl, 10 μL 500 mM TAPS pH 8.5, and water to a total volume of 50 μL. The solution was incubated at 95 °C on a heat block, slowly cooled over 1 h to the indicated temperature, and 1 μL 4M aqueous $NaBH_3CN$ was added. The reaction was allowed to proceed for the indicated time at the appropriate temperature, and then passed through a Princeton Separation sephadex minicolumn (eluted into 50 μL pure H_2O). The resulting sample was then evaporated to dryness in a speedvac concentrator, resuspended in 15 μL 1:1 TBE:formamide, heated at 95 °C for 15 min, and then separated by denaturing PAGE on a 15% (Figure 1) or 10% (Figures 2 and 3) TBE/urea polyacrylamide gel. All staining of PNA-DNA conjugates was done with ethidium bromide using standard DNA staining procedures. Only DNA-linked products are therefore visualized. Substantial readthrough product (corresponding to full-length polymerization through an obligate mismatched codon) is observed when reactions are allowed to proceed for much longer times or at low temperatures. The conditions reported in each Figure are optimized to generate significant full-length product with a minimum of readthrough product (Figures 1-3).

Control DNA-templated polymerization in the presence of all nine gvvt PNA tetramers. To verify that full-length DNA-templated polymerization (to generate a template-linked PNA 40-mer after ten successive coupling reactions) is not significantly inhibited by the presence of all nine gvvt PNA aldehyde tetramers, we compared side-by-side reactions under conditions listed in Figure 3 using (*i*) the DNA template in lines 1-2 of Figure 3 and (*ii*) a control reaction containing 16 μM of the gact PNA aldehyde and 2 μM of each of the other eight PNA aldehydes of sequence gvvt (32 μM total PNA aldehyde). The results are shown to the right and indicate that the generation of full-length product is not inhibited by the presence of all nine PNA aldehydes and does not lead to an increase in truncated product formation.



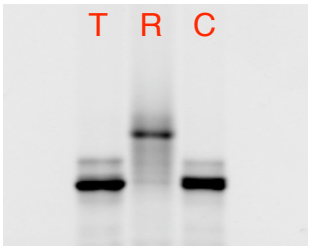
T = template only

gact = reaction with 0.4 μM template + 16 μM gact PNA aldehyde (**1**)

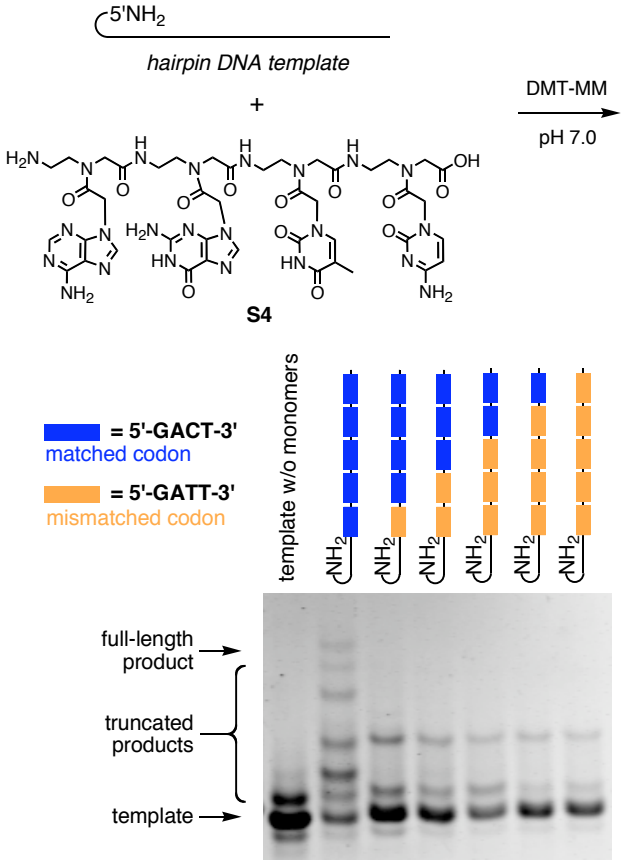
gvvt = reaction with 0.4 μM template + 16 μM gact PNA aldehyde (**1**) + 2 μM of each of the eight other PNA aldehydes of sequence gvvt (32 μM total PNA aldehyde)

Control reaction without sodium cyanoborohydride. In order to verify that observed oligomeric products contained secondary amine linkages between PNA tetramers (as opposed to unreduced imine linkages), we performed side-by-side oligomerization reactions with or without $NaBH_3CN$. The reaction shown in lane 2 of Figure 2 was repeated as described in the Figure 2 caption and above, with

the exception that the reaction time at room temperature in the presence of $NaBH_3CN$ was 30 min instead of 15 min. A second reaction was set up with the same conditions, except that no $NaBH_3CN$ was added to the reaction after annealing. Workup of the two reactions was performed as described above, and the products were analyzed by 10% denaturing PAGE. The absence of any observed coupling in the absence of $NaBH_3CN$ (see the Figure at right) verifies the requirement of $NaBH_3CN$ for product formation.



DNA-templated oligomerization by amine acylation. A PNA tetramer of sequence NH_2-agtc-COOH (**S4**) was synthesized on PAM resin by t-Boc chemistry. Loading of the C-terminal Boc-c-OH PNA monomer onto the resin was performed according the Applied Biosystems technical protocol *Manual esterification of hydroxymethyl-functionalized supports*. The basic PNA peptide synthesis protocol was adapted from Koch, T. *PNA oligomer synthesis by Boc chemistry*, in Peptide Nucleic Acids: Protocols and Applications, Nielsen and Egholm eds., Horizon Scientific Press, 1999. Amine acylation conditions were adapted from Gartner, Z.J.; Kanan, M.W.; Liu, D.R. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 1796. After optimization, the following reaction conditions were found to give the best yields. 20 pmol 5-amino-terminated DNA template (in 1□L) and 400 pmol NH_2-agtc-COOH (in 1 □L) were mixed together with 12.5 □L 2M NaCl and 10 □L 500 mM MOPS pH 7.0. The solution was heated for 10 min at 80 °C, slowly cooled over 1 h to room temperature, and then 25 □L 100 mM DMT-MM was added. The reaction was allowed to proceed for 24 h at room temperature, after which it was subjected to gel filtration (Princeton Separation) and evaporated to dryness. The sample was then resuspended in 1:1 TBE:formamide, heated at 95 °C for 15 min, and then separated by denaturing PAGE on a 15% TBE/urea polyacrylamide gel. Gels were stained with ethidium bromide using standard procedure. This amine acylation polymerization reaction with NH_2-agtc-COOH was repeated using a number of different 5-amino-terminated hairpin templates with mismatched codons at different positions. The oligonucleotide sequences that correspond to the templates used in the figure to the right are:



Lane 1:
NCGCGAGCGTACGCTCGCGAGACTGACT
GACTGACTGACTTTTT
Lane 2:
NCGCGAGCGTACGCTCGCGAGACTGACT
GACTGACTGACTTTTT
Lane 3:
NCGCGAGCGTACGCTCGCGAGATTGACT
GACTGACTGACTTTTT
Lane 4:

NCGCGAGCGTACGCTCGCGAGATTGATTGACTGACTGACTTTTT
Lane 5: NCGCGAGCGTACGCTCGCGAGATTGATTGATTGACTGACTTTTT
Lane 6: NCGCGAGCGTACGCTCGCGAGATTGATTGATTGATTGACTTTTT
Lane 7: NCGCGAGCGTACGCTCGCGAGATTGATTGATTGATTGATTTTTT

Sample preparation and MALDI-TOF analysis of polymerization products.
The DNA-templated polymerization reactions shown in lane 2 of Figure 1 and lane 2 of Figure 2 were scaled up 5-fold. Each of these samples was loaded onto preparative 10% denaturing polyacrylamide TBE/urea gels, and the product bands excised. The excised polyacrylamide was homogenized, soaked in TBE buffer overnight, and filtered. Due to the difficulty of ethanol precipitating PNA-DNA conjugates, the gel purification eluates were concentrated, and the resulting solution passed through a sephadex gel filtration column (Princeton Separation). Samples were resuspended in 49 μL 1x RQ1 DNase I buffer (Promega), and 1 μL RQ1 DNase I (Promega) was added. Following incubation at 37 °C for 1 h, 5 μL 10x RQ1 stop buffer was added, and the samples were heated at 65 °C for 30 min (to denature the enzyme), passed through Princeton Separation sephadex columns, and evaporated to dryness. The samples were resuspended in 5 μL pure H_2O, and submitted for MALDI-TOF mass spectrometry (with a sinapinic acid matrix) to the Harvard University Microchemistry Facility. In all cases, the observed masses of polymerization products and fragmentation species analyzed in this manner corresponded to structures ending in NH_2 (or OH) groups at the C-terminal (formerly aldehyde) end of the PNA products. Assuming this form of product species, for the full-length product observed in Figure 1 (lane 2), expected mass = 5366; observed mass = 5375 ± 15. For the full-length product observed in Figure 2 (lane 2), expected mass = 10719; observed mass = 10729 ± 30.

Comparison with previous DNA-templated reductive amination oligomerizations (Ref. 3b)
Lynn and co-workers describe the efficient oligomerization of DNA analogs by reductive amination in Ref. 3b and show that oligomerization proceeds by a step-growth mechanism. It is tempting to speculate that the DNA-template PNA oligomerizations and polymerizations described in this work may, in contrast, proceed via site-selective initiation and chain growth polymerization. We are primarily interested in those oligomers and polymers that are covalently linked to their amplifiable DNA templates and therefore have focused our analysis on these products. Short PNA tetramers and octamers not conjugated to DNA templates are often not visible on stained PAGE gels. Therefore, we cannot rigorously conclude at this point if the initiation of polymerization is exclusively site-specific, or if some coupling events take place at multiple locations on the template. Nevertheless, the (*i*) the similar efficiency of long polymerizations regardless of the placement of mismatched codons in the template, (*ii*) the presumed difficulty of any PNA octamer or longer oligomer (resulting from one or more coupling events) dissociating appreciably from the DNA template, and (*iii*) the high degree of observed sequence specificity are *consistent* with a site-specifically initiated chain-growth polymerization mechanism even though we have not rigorously precluded the possibility that some of the monomers couple together while annealed on the template but before becoming covalently linked to the template.