Supporting Information

Effects of Template Sequence and Secondary Structure on DNA-Templated Reactivity

Thomas M. Snyder, Brian N. Tse, and David R. Liu*
Howard Hughes Medical Institute and the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

DNA Sequences and Structures

The sequences and structures of functionalized DNA oligonucleotides used in this work are as follows. The template oligonucleotides (**1-8**, **15-26**) were synthesized using 3'-(6-Fluorescein) CPG and 5'-Amino Modifier 5 Phosphoramidite from Glen Research. The reagent oligonucleotides (**9-14**) were synthesized using 3'-Amino-Modifier C7 CPG 500 from Glen Research.

1 – 5'-(NH_2-)CCCTGTACACTGGGTAAACTTCATGGAATCATTCCATGCTACCCACTC-Fluorescein

2 – 5'-(NH_2-)CCCTGTACACTGGGTAAACTTGCATGTATCATTCCATGCTACCCACTC-Fluorescein

3 – 5'-(NH_2-)CCCTGTACACTCTAGGAACTTACATGGATCATTCCATGCTACCCACTC-Fluorescein

4 – 5'-(NH_2-)CCCTGTACACTCTAGGAACTTCATGGTATCATTCCATGCTACCCACTC-Fluorescein

5 – 5'-(NH_2-)CCCTGTACACTCTAGGAACTTTATGGCATCATTCCATGCTACCCACTC-Fluorescein

6 – 5'-(NH_2-)CCCTGTACACTCTAGGAACTTCATGCTATCATTCCATGCTACCCACTC-Fluorescein

7 – 5'-(NH_2-)CCCTGTACACTCTAGGAACTTCACTCAATCATTCCATGCTACCCACTC-Fluorescein

8 – 5'-(NH_2-)CCCTGTACACTCAACCAACTTCACTCAATCATTCCATGCTACCCACTC-Fluorescein

9 – 5'-GTGTACAGGG-NH_2

10 – 5'-TAGCATGGAAT-NH_2

11 – 5’-TAGCATGGAATAGGG-NH_2

12 – 5’-TAGTCAGTCATAGGG-NH_2

13 – 5’-AGCATGGAATAGGG-NH_2

14 – 5’-GTAGCATGGAATAGGG-NH_2

15 – 5’-(NH_2-)CCCT[A_{26}]ATTCCATGCTACCCACTC-Fluorescein

16-**22** are templates, similar to the poly-A template **15**, but with random mixes used in place of A_{26}. The general sequence is: 5’-(NH_2-) CCCT[X_{26}]ATTCCATGCTACCCACTC-Fluorescein where X is a nucleotide mix code: M (A and C) for **16**, K (G and T) for **17**, Y (C and T) for **18**, R (A and G) for **19**, W (A and T) for **20**, S (C and G) for **21**, N (A, C, G, and T) for **22**.

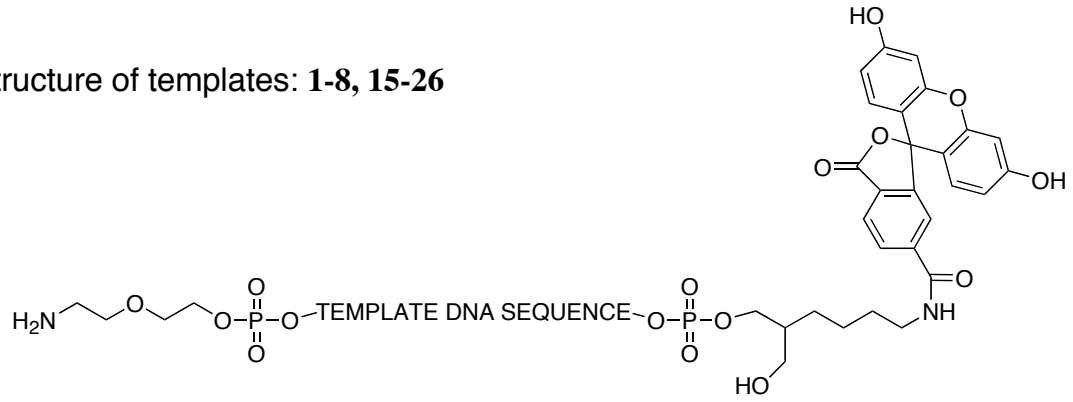
23 - 5’-(NH_2-)CCCTGTACACTCAACCAACTTCAGTTGATCATTCCATGCTACCCACTC-Fluorescein

24 - 5’-(NH_2-)CCCTGTACACGATCCAAACTTGTTTGGATCATTCCATGCTACCCACTC-Fluorescein

25 - 5’-(NH_2-)CCCTGTACACTCAACCAACTTAGGTACATCATTCCATGCTACCCACTC-Fluorescein

26 - 5’-(NH_2-)CCCTGTACACTGATCCAAGATCAGTGTACCATTCCATGCTACCCACTC-Fluorescein

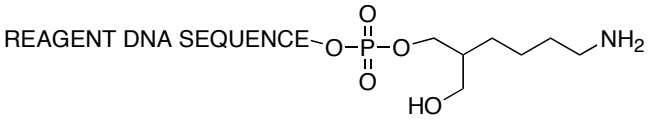
Structure of templates: **1-8, 15-26**



from 5'-Amino Modifier 5 phosphoramidite

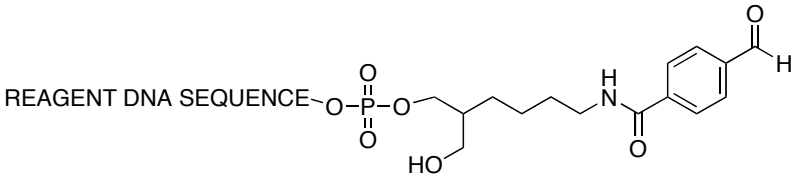
from 3'-(6-Fluorescein) CPG

Structure of reagents (following DNA synthesis): **9, 10, 11, 12, 13, 14**



from 3'-Amino-Modifier C7 CPG

Structure of reagents (for reductive amination): **9a, 10a, 11a, 12a, 13a, 14a**



Structure of reagents (for amine acylation): **9b, 10b, 11b, 12b, 13b, 14b**

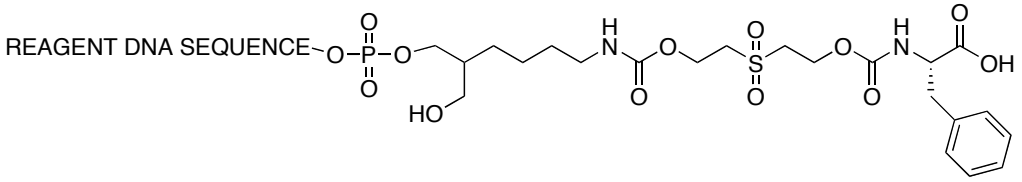


Figure S1. Chemical structures of the DNA templates and reagents used in this study.

Additional Computational Modeling of Templates and Reagents. To verify our sequence analysis with OMP, we folded each of the templates **1** to **8** using the MFOLD DNA QuikFold server (http://frontend.bioinfo.rpi.edu/applications/hybrid/quikfold.php) with the specific conditions of 25 °C, 1 M [Na^+], 0 M [Mg^{2+}]. The reported free energies for the most stable structures are given in Table S1 for both OMP and MFOLD. MFOLD also outputs predicted structures and these were compared to those generated by OMP (Figure 2). The predicted structures are identical except that MFOLD excludes some of the AT pairs at the ends of stems, specifically those in templates **1**, **2**, **5**, and **8**. Also, for templates **3** through **6**, MFOLD predicts an additional structural element containing a three-base stem forming at the 5' primer-binding site which may account for the increase in energies of these four templates, in particular, over the OMP predicted values.

	Template							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
OMP	-10.1	-8.57	-7.51	-5.79	-3.87	-2.87	-1.58	+0.11
MFOLD	-9.95	-8.83	-8.41	-6.24	-4.82	-3.73	-2.87	-0.44

Table S1. Free energies of folding for the eight templates (in kcal/mole) as estimated by OMP and by MFOLD.

MFOLD was also used to analyze the folding of the thirty 5'-terminal bases in templates **15** and **23-26**. For **15**, MFOLD reports that there is "no structure possible" which is not unexpected given 26 of the 30 bases are adenine. For **23**, MFOLD reports a slightly varied structure with a smaller loop size but a similar energy (-4.16 kcal/mole for MFOLD versus -3.38 kcal/mole). The structures for **24**, **25**, and **26** are the same in MFOLD and OMP and the predicted energies are within 0.5 kcal/mole for each case. Overall, while the exact values vary between the two software packages, the predicted structures and general energetic trends are reasonably conserved between OMP and MFOLD.

We also used NUPACK (www.nupack.org) to model the hybridization of reagent **9** and **10** to each of the eight templates and compared these results to those from OMP. For the NUPACK analysis, 25 °C was used as the melting temperature with 100 nM template and 150 nM reagent as for OMP. NUPACK does not allow specific parameters for salt which may cause

some variation from the OMP conditions (1 M [Na^+], 0 M [Mg^{2+}]). In both cases, the control reagent **9** is bound to all templates at greater than 99%. For reagent **10**, the values between the two software packages vary (Table S2) but the general trend is conserved, with the more structured templates being less bound by reagent than the less structured templates.

	Template							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
OMP	0.1	1.3	7.5	21.4	71.4	92.8	99.0	99.5
NUPACK	1.6	22.0	25.3	64.4	88.4	97.4	99.0	99.6

Table S2. Predicted hybridization efficiencies between reagent **10** and the templates **1**-**8** using OMP and NUPACK. Reported values represent the percentage of templates at equilibrium bound by reagent. Analysis was performed with 100 nM template, 150 nM reagent, at 25 °C in both cases.

Reactivity of Templates Using the End-of-Helix Architecture and Amine Acylation. We performed experiments to study amine acylation using the end-of-helix architecture similar to the reported experiments involving reductive amination. A control reagent (**9b**) that binds at the 5'-primer binding site was first used to measure the maximal yield when hybridization efficiency and separation between functional groups are not affecting reactivity. As seen in Table S3, these reactions (**1-8** with **9b** in 0.1 M MES buffer pH 6.0, 1 M NaCl, 24 mM sNHS, and 32 mM EDC, 25 °C for 8 h) resulted in good yields of 72-78 % as determined by denaturing PAGE analysis and fluorophore densitometry.

We then reacted an 11-base reagent (**10b**) that anneals 30 bases away from the end of the template under amine acylation conditions. (Table S3) Templates **1** and **2**, with the most internal secondary structure, react to give 5 % and 9 % product yields, respectively. The reactivity of template **3** is significantly higher, though still modest, resulting in 27 % yield of product. As we observed for the reductive amination reaction, product yields increase for these first three templates as the amount of internal template secondary structure was reduced.

Once again a new trend emerged when measuring the reactivity of templates **4**-**8**. Templates **4**-**6** react at 14-18 % yield, only about half of the yield for template **3**. Template **7** reacts at 6 % yield while template **8** reacts at just 4 % yield, worse than all other templates as we

also observed with reductive amination. The yield for template **8** is twenty-fold below the yield observed for the control reagent **9b** with template **8**. Just as with the earlier end-of-helix reactivity results for reductive amination, the structure of the 30-base intervening region of the template appears to be strongly affecting the reactivity for these templates, and the templates with very little internal structure are virtually unreactive.

	Template							
Reagent	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
9b	76	74	78	75	73	77	72	76
10b	5	9	27	18	16	14	6	4
ΔG (kcal/mol)	-10.1	-8.57	-7.51	-5.79	-3.87	-2.87	-1.58	+0.11

Table S3. Product yields (in %) for templates **1-8** reacting with reagents using the end-of-helix architecture and amine acylation. Reactions were performed with 150 nM reagent, 100 nM template in 0.1 M MES buffer, pH 6.0, 1.0 M NaCl, 32 mM EDC, and 24 mM sNHS for 8 h at 25 °C. The folding energies of the eight templates are listed below the product yields.

Effects of Reagent Length on the Amine Acylation Reaction. Similar to our experiments with reductive amination shown in Figure 4, we performed reactions with varied reagent lengths to determine the effect of hybridization on amine acylation. The shorter, 10-base amine acylation reagent **13b** reduced product yields for highly and moderately structured templates **1-6**, while the longer, 12-base amine acylation reagent **14b** increased product yields for these same templates. (Table S4) For the most highly structured template (**1**), the reaction yields for this 10-, 11-, and 12-base reagent series (**13b**, **11b**, **14b**) were 4 %, 7 %, and 18 % yield respectively, with increasing yields as reagent length increased. For template **2**, which is also very structured, the yields were 12 % for the 10-base reagent (**13b**), 19 % for the 11-base reagent (**11b**), and 31 % for the 12-base reagent (**14b**). These results show a clear trend connecting template-reagent hybridization to the efficiency of templated reactions.

However, similar to our observations with the reductive amination reaction, neither the shorter (**13b**) nor the longer (**14b**) reagent significantly altered the reactivity of highly unstructured templates **7** and **8** (Table S4). For all three reagent lengths, the amine acylation yields for template **7** ranged over a narrow range of 44-46 % yield. For template **8**, the yields with reagents **13b**, **11b**, and **14b** ranged from 32-37 % yield. These results further indicate that

an unusually low amount of internal structure among intervening template oligonucleotides impairs DNA-templated reactivity in a manner that is uncoupled from template-reagent hybridization.

	Template							
Reagent	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
13b	4	12	28	27	46	60	46	32
11b	7	19	41	40	56	64	46	37
14b	18	31	51	58	49	59	44	32

Table S4. Product yields (in %) for templates **1-8** reacting with the reagents listed in the omega architecture. Reagents **13b, 11b**, and **14b** contain 10-, 11-, and 12-base regions, respectively, that complement codon 3. As reagent length increases, the reactivity for highly structured templates (**1-4**) increases substantially, while the reactivity of the unstructured templates (**7-8**) is not significantly affected.

Complete References **24-26**:

(24) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C. *Science* **2001**, *291*, 1304-51.

(25) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.;

Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendl, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J. F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M. *Nature* **2001**, *409*, 860-921.

(26) Kapranov, P.; Cheng, J.; Dike, S.; Nix, D. A.; Duttagupta, R.; Willingham, A. T.; Stadler, P. F.; Hertel, J.; Hackermuller, J.; Hofacker, I. L.; Bell, I.; Cheung, E.; Drenkow, J.; Dumais, E.; Patel, S.; Helt, G.; Ganesh, M.; Ghosh, S.; Piccolboni, A.; Sementchenko, V.; Tammana, H.; Gingeras, T. R. *Science* **2007**, *316*, 1484-8.