J|A|C|S ARTICLES

Published on Web 03/04/2010

Expanding the Chemistry of DNA for in Vitro Selection

Jonathan D. Vaught,[†,‡] Chris Bock,[‡] Jeff Carter,[‡] Tim Fitzwater,[‡] Matt Otis,[‡] Dan Schneider,[‡] Justin Rolando,[†] Sheela Waugh,[‡] Sheri K. Wilcox,[‡] and Bruce E. Eaton*[,†]

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, and SomaLogic, Inc., 2945 Wilderness Place, Boulder, Colorado 80301

Received September 22, 2009; E-mail: bruce.eaton@colorado.edu

Abstract: Six new 5-position modified dUTP derivatives connected by a unique amide linkage were synthesized and tested for compatibility with the enzymatic steps of in vitro selection. Six commercially available DNA polymerases were tested for their ability to efficiently incorporate each of these dUTP derivatives during PCR. It was not possible to perform PCR under standard conditions using any of the modified dUTP derivatives studied. In contrast, primer extension reactions of random templates, as well as defined sequence templates, were successful. KOD XL and D. Vent DNA polymerases were found to be the most efficient at synthesizing full-length primer extension product, with all of the dUTP derivatives tested giving yields similar to those obtained with TTP. Several of these modified dUTPs were then used in an in vitro selection experiment comparing the use of modified dUTP derivatives with TTP for selecting aptamers to a protein target (necrosis factor receptor superfamily member 9, TNFRSF9) that had previously been found to be refractory to in vitro selection using DNA. Remarkably, selections employing modified DNA libraries resulted in the first successful isolation of DNA aptamers able to bind TNFRSF9 with high affinity.

Introduction

RNA and DNA in vitro selection methods[1−3] have yielded specific sequences that function as biopolymer ligands (aptamers) and catalysts for a wide range of applications in therapeutics[4,5] and diagnostics.[6−9] Both DNA and RNA aptamers in their native form have been selected to specifically bind proteins and small molecules.[10−13] In addition, a growing list of organic reactions are catalyzed by specific nucleic acid sequences. These include acyl transfers,[14−18] aldol condensation,[19] amide bond formation,[20] Diels−Alder cycloaddition,[21−25] Michael additions,[26] nucleophilic substitutions,[27,28] porphyin metalations,[29] and enzyme cofactor synthesis.[30,31]

There are limited data on the use of modified nucleotides in either DNA or RNA to select modified sequences with functions beyond that of native DNA and RNA. In most of biology, proteins are the predominant catalysts (e.g., enzymes) and affinity molecules (e.g., antibodies), although there are notable exceptions such as the ribosome peptidyl transferase, which is catalyzed by RNA.[48] A common explanation for the functional dichotomy between nucleic acids and proteins is that amino

[†] University of Colorado.
[‡] SomaLogic, Inc.

(1) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
(2) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
(3) Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850–852.
(4) Lee, J. F.; Stovall, G. M.; Ellington, A. D. *Curr. Opin. Chem. Biol.* **2006**, *10*, 282–289.
(5) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555.
(6) Chu, T. C.; Shieh, F.; Lavery, L. A.; Levy, M.; Richards-Kortum, R.; Korgel, B. A.; Ellington, A. D. *Biosens. Bioelectron.* **2006**, *21*, 1859–1866.
(7) Collett, J. R.; Cho, E. J.; Lee, J. F.; Levy, M.; Hood, A. J.; Wan, C.; Ellington, A. D. *Anal. Biochem.* **2005**, *338*, 113–123.
(8) Jhaveri, S.; Rajendran, M.; Ellington, A. D. *Nat. Biotechnol.* **2000**, *18*, 1293–1297.
(9) Brody, E. N.; Willis, M. C.; Smith, J. D.; Jayasena, S.; Zichi, D.; Gold, L. *Mol. Diagn.* **1999**, *4*, 381–388.
(10) Osborne, S. E.; Ellington, A. D. *Chem. Rev.* **1997**, *97*, 349–370.
(11) Famulok, M.; Szostak, J. W. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 979–988.
(12) Eaton, B. E. *Curr. Opin. Chem. Biol.* **1997**, *1*, 10–16.
(13) Gold, L.; Polisky, B.; Uhlenbeck, O.; Yarus, M. *Annu. Rev. Biochem.* **1995**, *64*, 763–797.
(14) Lohse, P. A.; Szostak, J. W. *Nature* **1996**, *381*, 442–444.
(15) Suga, H.; Lohse, P. A.; Szostak, J. W. *J. Am. Chem. Soc.* **1998**, *120*, 1151–1156.
(16) Lee, N.; Bessho, Y.; Wei, K.; Szostak, J. W.; Suga, H. *Nat. Struct. Biol.* **2000**, *7*, 28–33.
(17) Illangasekare, M.; Sanchez, G.; Nickles, T.; Yarus, M. *Science* **1995**, *267*, 643–647.
(18) Kumar, R. K.; Yarus, M. *Biochemistry* **2001**, *40*, 6998–7004.
(19) Fusz, S.; Eisenfuhr, A.; Srivatsan, S. G.; Heckel, A.; Famulok, M. *Chem. Biol.* **2005**, *12*, 941–950.
(20) Wiegand, T. W.; Janssen, R. C.; Eaton, B. E. *Chem. Biol.* **1997**, *4*, 675–683.
(21) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *Nature* **1997**, *389*, 54–57.
(22) Seelig, B.; Jaschke, A. *Chem. Biol.* **1999**, *6*, 167–176.
(23) Tarasow, T. M.; Tarasow, S. L.; Tu, C.; Kellogg, E.; Eaton, B. E. *J. Am. Chem. Soc.* **1999**, *121*, 3614–3617.
(24) Stuhlmann, F.; Jaschke, A. *J. Am. Chem. Soc.* **2002**, *124*, 3238–3244.
(25) Tarasow, T. M.; Tarasow, S. L.; Eaton, B. E. *J. Am. Chem. Soc.* **2000**, *122*, 1015–1021.
(26) Sengle, G.; Eisenfuhr, A.; Arora, P. S.; Nowick, J. S.; Famulok, M. *Chem. Biol.* **2001**, *8*, 459–473.
(27) Wecker, M.; Smith, D.; Gold, L. *RNA* **1996**, *2*, 982–994.
(28) Wilson, C.; Szostak, J. W. *Nature* **1995**, *374*, 777–782.
(29) Li, Y. F.; Sen, D. *Nat. Struct. Biol.* **1996**, *3*, 743–747.
(30) Huang, F. Q.; Bugg, C. W.; Yarus, M. *Biochemistry* **2000**, *39*, 15548–15555.
(31) Jadhav, V. R.; Yarus, M. *Biochimie* **2002**, *84*, 877–888.

acids have a wider array of chemical functionality than nucleotides. Few examples have been reported where the affinity or catalytic activity of selected aptamers or nucleic acid catalysts could be enhanced by adding chemical diversity to nucleobases. Most of these experiments have focused on RNA, and the results have been mixed. It remains an open scientific question as to which biopolymer, RNA or DNA, provides the better framework (for catalysis or aptamers) and if chemical modification of the nucleobases can be generally beneficial. Two important challenges for the field are the synthesis of modified nucleotides and their enzymatic incorporation into nucleic acids for in vitro selection. Herein we describe the possibilities for creating the same type of 5-position uridine modification chemistry in DNA that has been successful in RNA selections.

The majority of data reported on modified aptamers is from RNA modified at the 2′-position of ribose, a modification used to increase the in vivo stability of RNA.[32] Modified DNA has been used to select both catalysts and aptamers. Elegant landmark examples have been reported where base-modified deoxyuridine and deoxyadenosine were incorporated into DNA catalysts for phosphodiester bond cleavage that, unlike their unmodified RNA counterparts, did not require divalent metal cations.[35] Examples of modified DNA aptamers are also known. Recently, the selection of an enantioselective thalidomide aptamer was reported that contained a base-modified deoxyuridine, which may prove to be useful as an enantiospecific bioanalytic reagent.[40] However, not all modified chemistries for nucleic acids have proven beneficial. In 1994, a thrombin aptamer was selected using a modified deoxyuridine.[41] In comparison to a previously known unmodified DNA aptamer, the affinity of the modified aptamer for thrombin was significantly worse. It appears that the functional group of the modified DNA, and perhaps the chemistry of the linkage to the nucleobase, may be important in determining if a functional group has any advantage in an aptamer for high affinity binding to a protein.[42]

To be useful for in vitro selection experiments, modified triphosphates must serve as substrates for polymerases used to replicate the oligonucleotide under investigation. In the case of RNA, the kinetics of enzymatic incorporation by T7 RNA polymerase for a series of 5-position carboxamid-modified UTP derivatives has been previously reported.[38] Two of these UTP derivatives have been used successfully for in vitro selection of RNA catalysis for organic reactions that include nontemplated amide bond formation,[20] a Diels−Alder cycloaddition with a low reactivity acyclic diene,[21] and urea bond formation with stereoselectivity for tripeptide substrates.[39] In addition to the modified deoxynucleotides used to select a thalidomide aptamer, thrombin aptamer, and a catalyst for phosphodiester bond cleavage, several groups have synthesized modified deoxynucleotides using conventional alkyne linkage chemistries to the nucleobase. These are proven substrates for certain DNA polymerases but have yet to be reported as successful for in vitro selection experiments for aptamers to protein targets.[33–37]

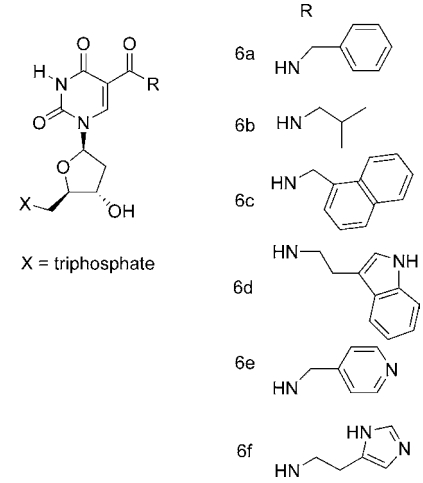


Figure 1. New dUTP derivatives prepared as described in Scheme 1.

To better understand what modifications in DNA could yield improved aptamers and catalysts, more chemical examples are needed. The restricted rotation of the amide bond and its ability to serve as both hydrogen bond acceptor and donor appeared to be attractive attributes for the linker since new H-bonding contacts may be possible with protein targets. Furthermore, minimizing the number of rotable bonds between the nucleobase and the pendant modification seemed desirable to allow for precise orientation at the aptamer protein interface.

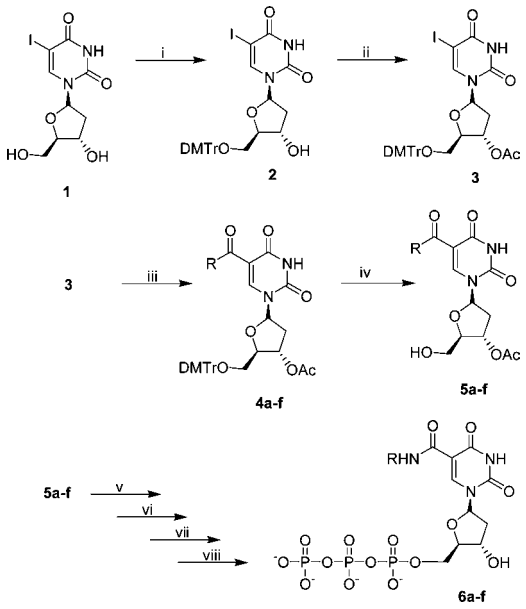
To efficiently survey new DNA chemistries required a streamlined chemical modification synthetic procedure. Herein we describe adaptation of our previously reported methods for preparing modified uridines to the synthesis of 5-position modified 2′-deoxyuridines, preparing the corresponding dUTP derivatives, testing these as substrates with six DNA polymerases, and establishing viable in vitro selection methods. Finally, we will describe using these dUTP derivatives to form DNA libraries and how these libraries were successful for in vitro selection to yield a high affinity aptamer sequence to a challenging human protein target TNFRSF9.

Results and Discussion

Synthesis of 5-Position Carboxamid-2′-deoxyuridines and Triphosphates. In order to expand the repertoire of DNA, it was desirable to attach functional groups to the 5-position of 2′-deoxyuridine by an amide linkage that may provide new H-bonding motifs and concomitant 3D structures. Further, it was of interest to synthesize a range of hydrophobic groups, both aromatic and aliphatic, to increase the hydrophobic content (Figure 1) of DNA to potentially mimic the types of side chains found in the binding domains of antibodies. These functional groups were added to 2′-deoxyuridine in a single step by palladium-catalyzed carboxyamidation of 5-iodo-2′-deoxyuridine under mild conditions (Scheme 1). The palladium-catalyzed 5-position modification of uridines can also be accomplished

(32) Aurup, H.; Williams, D. M.; Eckstein, F. *Biochemistry* **1992**, *31*, 9636–9641.
(33) Sakthivel, K.; Barbas, C. F. *Angew. Chem., Int. Ed.* **1998**, *37*, 2872–2875.
(34) Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071–15082.
(35) Perrin, D. M.; Garestier, T.; Helene, C. *Nucleosides, Nucleotides Nucleic Acids* **1999**, *18*, 377–391.
(36) Gourlain, T.; Sidorov, A.; Mignet, N.; Thorpe, S. J.; Lee, S. E.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1898–1905.
(37) Benner, S. A. *Acc. Chem. Res.* **2004**, *37*, 784–797.
(38) Vaught, J. D.; Dewey, T.; Eaton, B. E. *J. Am. Chem. Soc.* **2004**, *126*, 11231–11237.
(39) Nieuwlandt, D.; West, M.; Cheng, X. Q.; Kirshenheuter, G.; Eaton, B. E. *Chembiochem* **2003**, *4*, 651–654.
(40) Shoji, A.; Kuwahara, M.; Ozaki, H.; Sawai, H. *J. Am. Chem. Soc.* **2007**, *129*, 1456–1464.
(41) Latham, J. A.; Johnson, R.; Toole, J. J. *Nucleic Acids Res.* **1994**, *22*, 2817–2822.
(42) Tarasow, T. M.; Eaton, B. E. *Biopolymers* **1998**, *48*, 29–37.

Scheme 1. Synthesis of **6a**−**f**[a]



[a] Conditions: (i) 1.1 equiv of DMTrCl, pyridine, rt, 16 h; (ii) 10 equiv of Ac_2O, pyridine, rt, 16 h; (iii) 10 mol % of $Pd(PPh_3)_4$, 5 equiv of TEA, 3 equiv of RNH_2, DMA/THF, 50 psi CO, 70 °C, 48−72 h; (iv) 1 equiv of TCA, MeOH/CH_2Cl_2 1:1, rt, 2−4 h; (v) 1.1 equiv of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one, DMF/pyridine/dioxane 1:1:4, rt, 10 min; (vi) 1.5 equiv of tributylammoniun pyrophosphate, tributylamine, DMF, rt, 10 min; (vii) 0.02 M I_2, THF/pyridine/H_2O 70:20:10, rt. 15 min; (viii) concd NH_4OH, 1 h.

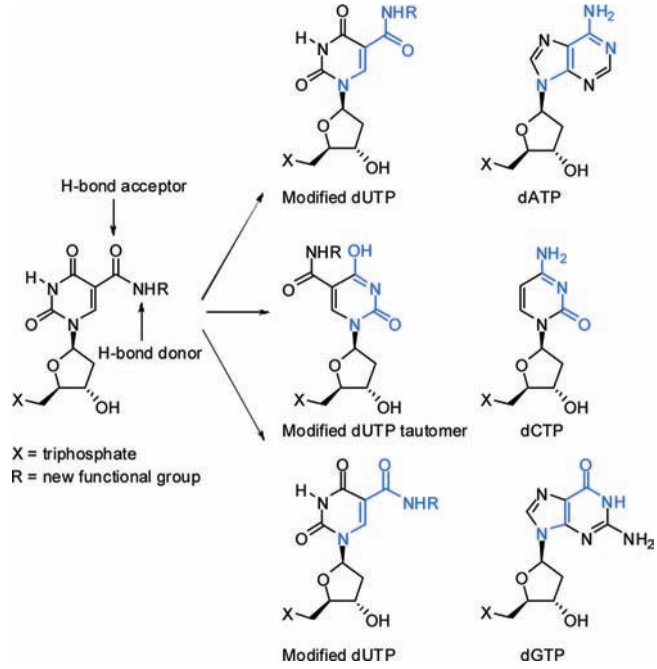


Figure 2. Some possible orientations of 5-position modifications on dUTP that lead to new H-bonding motifs. Atoms and bonds shown in blue indicate highly similar regions where a modified dUTP could present itself almost identically to each of the native nucleotides.

using stannanes to provide a ketone linkage to the functional group.[43] The utility of stannanes for incorporation of functional groups is limited, however, because stannanes are difficult to prepare. An alternative palladium-catalyzed reaction is the carbonylative coupling using amines.[44] In contrast to stannanes, numerous amines are readily available. However, when applying these 5-position carbonylative coupling methods to 2′-deoxyuridines, different nucleoside protection and coupling procedures must be followed to achieve good yields.

Starting with commercially available 5-iodo-2′-deoxyuridine, it was necessary to first protect the 5′-hydroxyl with a DMT group to give **1** followed by acetylation of the 3′-hydroxyl to give the differentially protected nucleoside **3**. Linkage of the various functional groups was accomplished by palladium ($Pd[PPh_3]_4$, 10 mol %)-catalyzed carbonylative coupling with primary amines and CO (50 psi) in DMF/THF at 70 °C, using 5 equiv to scavenge the HI produced in forming **4a**−**f**. Following purification of **4a**−**f** by column chromatography, the 5′-DMT protecting group was removed with trichloroacetic acid (1 equiv, in methanol/CH_2Cl_2, rt). With the 5-hydroxyl deprotected, triphosphate synthesis was accomplished using the Eckstein[45] triphosphate method to give **6a**−**f** in good to moderate (30−60%) yield.

While there are other chemistries available for introducing modifications to the 5′-position of 2′-deoxyuridine, this chemistry is unique in that the modifications are linked through an amide bond directly to the pyrimidine base at the C5-position, which is much less hydrophobic than alkyne linkers and hydrocarbon chains frequently used in other synthetic schemes for making base-modified nucleotides. Also, the amide bond introduces two new hydrogen-bonding contacts, a hydrogen bond donor and hydrogen bond acceptor (Figure 2). This amide is cross-conjugated with the pyrimidine ring which has significant electronic and structural effects on the pyrimidine ring system (Eaton, B. E. and Rolando, J., unpublished results) in addition to the effects of the appended functional group connected by the amide. For example, the tautomerization of dUTP to make the hydroxyl at C4 could be facilitated by the presence of the amide hydrogen forming an intermediate H-bond with the oxygen at C4. Another possibility is that, depending on the orientation of the amide with respect to the ring, the modified uridine base could display hydrogen-bonding patterns similar to those observed in Hoogsteen binding, where the bases are dA, dG, or dC. In the context of a folded ssDNA, alternative base pairings (e.g., duplex, triplex, etc.) could be advantageous, providing alternative hydrogen-bonding motifs that stabilize unique three-dimensional structures. Another key feature of the C5-modification chemistry described herein is that the number of rotable bonds between the functional group and the base is minimized, making these types of connections to the modifications more like those found connecting the functional groups found in proteins.[42] Lastly, the types of functional groups appended are similar to some of the most common critical residues in antibody CDR regions, mainly aromatic groups with varying degrees of hydrophobicity and polarizability.

DNA Polymerase Screening for Modified DNA Synthesis by Primer Extension. In order to determine whether these new modified nucleotides could be incorporated enzymatically into DNA, we tested the ability of commercially available DNA polymerases to incorporate the dUTP derivatives **6a**−**f** in primer extension reactions. It should be noted that several groups have previously reported modified DNA synthesis using 5-position modified dUTP derivatives with various DNA polymerases during PCR and primer extension reactions. However, none of

(43) Dewey, T. M.; Mundt, A. A.; Crouch, G. J.; Zyzniewski, M. C.; Eaton, B. E. *J. Am. Chem. Soc.* **1995**, *117*, 8474–8475.
(44) Dewey, T. M.; Zyzniewski, M. C.; Eaton, B. E. *Nucleosides Nucleotides* **1996**, *15*, 1611–1617.
(45) Ludwig, J.; Eckstein, F. *J. Org. Chem.* **1989**, *54*, 631–635.

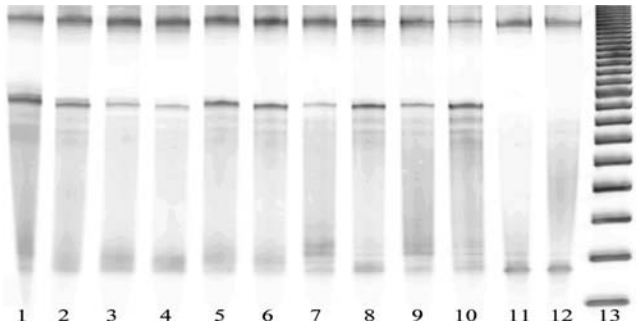


Figure 3. Fluorescent image of primer extension products from a 92 nucleotide biotinylated ssDNA template using TTP or **6a**, and KOD XL (lanes 1 and 2), *Pfu* (exo-) (lanes 3 and 4), D. Vent (exo-) (lanes 5 and 6), *Tth* (lanes 7 and 8), or *Taq* (lanes 9 and 10). All reactions were incubated with streptavidin before gel loading. In lanes 1−10, odd numbered lanes are with **6a** and even numbered lanes are with TTP. Lanes 11 and 12 are control lanes, with no TTP and no enzyme, respectively. Lane 13 is a dsDNA ladder of 10 base increments starting with 20 bases.

the modifications used in these previously reported studies contained an amide linkage connected to the 5-position. In addition, longer linkers and even specific linker chemistries had been proposed[33,46] to be required to allow 5-position modified dUTP analogues to work as substrates for DNA polymerases. Further, it was typically required to adjust the PCR buffer and primer extension conditions to get useful yields of full-length DNA. These previous reports on the enzymatic incorporation of C5-position modified deoxynucleotide triphosphates made us concerned that the new triphosphates we had prepared would be only poorly incorporated into DNA, if at all.

Two groups of three polymerases each from Family A and Family B were screened using the dUTP analogue **6a** in addition to dATP, dCTP, and dGTP. Family A DNA polymerases included *Thermus aquaticus* (Taq), *Thermus thermohilus* (Tth), and the Klenow fragment (exo-). Family B DNA polymerases included *Pyroccus furiosus* (*Pfu*, exo-), *Pyrococcus* species GB-D (D. Vent), and *Thermococcus kodakaraensis* KOD1 (KOD XL).[47] Each polymerase was used in the commercially supplied buffer. To simulate what would be required for an in vitro selection, a 92-mer template was used containing a 40N random region flanked by fixed sequence regions for primer annealing. The template used for these studies was synthesized by automated DNA synthesis with biotin on the 5′-end so that the primer extension product modified DNA could be separated from the template strand by incubating the primer extension mixture with streptavidin subsequent to analysis by denaturing polyacrylamide gel electrophoresis. Visualization of DNA bands was achieved by staining the gel with SYBR Gold. A representative polyacrylamide gel of this screening reaction is shown in Figure 3. All polymerases except Klenow fragment (exo-) were able to incorporate **6a**, but there were significant differences between them and the utility of their incorporation was unclear. It appeared that Family B DNA polymerases were somewhat superior to Family A with regard to incorporation of this modified dUTP analogue.

(46) Lee, S. E.; Sidorov, A.; Gourlain, T.; Mignet, N.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565–1573.
(47) KOD XL is approximately 98% exo- and 2% native KOD. See: Nishioka, M.; Mizuguchi, H.; Fujiwara, S.; Komatsubara, S.; Kitabayashi, M.; Uemura, H.; Takagi, M.; Imanaka, T. *J. Biotechnol.* **2001**, *88*, 141–149.
(48) Polacek, N.; Mankin, A. S. *Crit. Rev. Biochem. Mol. Biol.* **2005**, *40*, 285–311.

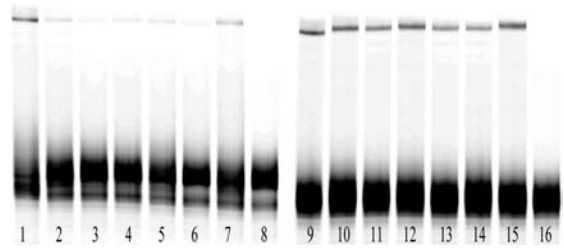


Figure 4. Fluorescent image of Cy5-labeled primer extension products from a 92 nucleotide ssDNA template with a 40N random region using TTP (lanes 1 and 9) or **6a**−**f** (lanes 2−7 and 10−15, respectively), and *Taq* (lanes 1−8), or D. Vent (exo-) (lanes 9−16). Lanes 8 and 16 are control lanes with no TTP or dUTP.

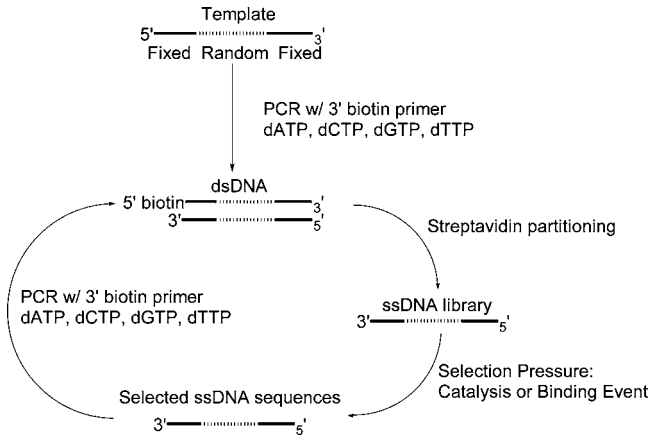
Table 1. Primer Extension Efficiencies of Various DNA Polymerases with **6a**−**f** (Percentages Obtained from the Ratio of the Amount of Full-Length Product from Primer Extensions with **6a**−**f** to the Amount of Product Obtained with TTP)

triphosphate	KOD XL	*Pfu* (exo-)	D. Vent (exo-)	*Tth*	*Taq*	KF (exo-)
TTP	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
6a	85.9%	39.3%	80.9%	16.0%	15.5%	1.1%
6b	83.6%	41.3%	64.1%	3.7%	6.4%	0.4%
6c	84.6%	28.4%	47.5%	6.1%	7.1%	0.2%
6d	88.0%	61.5%	100.0%	5.8%	8.9%	0.7%
6e	88.1%	34.4%	51.6%	16.5%	14.3%	1.5%
6f	90.3%	59.5%	99.4%	30.1%	32.1%	1.4%

It was next of interest to determine if **6a** was unique in its ability to be tolerated as a substrate by D. Vent polymerase and not well-tolerated by Taq polymerase. Each of the modified dUTP derivatives **6a**−**f** was tested with these two polymerases. In these primer extensions, a primer was used that had a 5′-Cy-5 fluorescent dye attached so that only the product DNA would be visualized during analysis by polyacrylamide gel electrophoresis (Figure 4). As seen previously for **6a**, Taq polymerase gave low to modest yield of full-length DNA product under the primer extension conditions used. In contrast, D. Vent gave good to excellent yields of full-length DNA product and appeared less sensitive to the type of modification attached to the dUTP. Primer extension experiments were then performed on all of the DNA polymerases of interest under the following conditions: DNA polymerase (1.5 U), dATP, dCTP, dGTP, TTP, or one of the modified dUTP derivatives **6a**−**f** (final concentration 200 μM each dNTP). Reactions were carried out at 70 °C for 30 min, except for KF reactions, which were carried out at 37 °C. Table 1 summarizes these primer extension screening results. The percent yield for each polymerase was obtained from the ratio of product from each primer extension reaction with **6a**−**f** relative to the amount of product from primer extension with TTP. Generally, Family B DNA polymerases gave greater yields of full-length DNA when incorporating **6a**−**f** than the Family A DNA polymerases. Surprisingly, KOD XL and D. Vent (exo-) were able to incorporate *all* of the modified dUTP derivatives **6a**−**f** with similar or better yields than with TTP. It will be noted that these yields were under screening conditions comparable to those that would be used during in vitro selection and should not be considered as relative Vmax etc.

Interestingly, the isobutyl-modified dUTP **6b** gave the lowest yields, as opposed to the uridines modified with larger aromatic groups such as indole or napthyl. As expected as a result of increased mass to charge ratio, all primer extension product DNA incorporating **6a**−**f** showed a decrease in electrophoretic mobility compared to DNA where TTP was incorporated.

Scheme 2. Generic DNA In Vitro Selection



PCR Amplification of 6a−f. A typical procedure for performing DNA in vitro selection to identify new catalysts and aptamers is shown in Scheme 2. The selection cycle begins with the chemical synthesis of a template DNA strand library (ca. 10^{14} sequences) composed of a random sequence flanked by fixed sequence regions required for PCR. PCR amplification of this template library in the presence of a 3′-primer containing a biotin at its 5′-end yields biotinylated dsDNA. Note that this step helps ensure that the sequences to go forward into selection are amplifiable. The dsDNA is then treated with streptavidin immobilized on beads, and the complement strand is removed from the beads by heating or raising the pH to generate a ssDNA library in solution. The desired selection pressure is then applied to the ssDNA for binding a target or performing a chemical reaction. The selected ssDNA sequences are then subjected to the same PCR procedure with the 3′-primer containing a biotin at its 5′-end to yield a dsDNA library enriched in the sequences suitable to perform the target binding or catalysis. In addition, a significant number of background sequences will also be carried forward so that the cycle must be repeated several times to isolate the best sequences as per the selection conditions.

DNA in vitro selection is especially attractive because, compared to RNA in vitro selection, only one enzymatic step (PCR) is required compared to three (transcription, reverse transcription, and PCR). Given the success of the primer extension studies using **6a−f**, it was of interest to determine if the Family A and Family B polymerases discussed above could function in PCR so that a DNA in vitro selection cycle could be demonstrated. Furthermore, since a pool of sequences was used in the primer extension experiments, it was unclear if incorporation of multiple 5-position modified 2′-deoxyuridines within a short sequence would result in lower yields of full-length DNA. A test sequence complementary to the sequence shown in Figure 5 was designed and prepared by automated DNA synthesis. Real-time PCR using **6a−f** was performed using the complement to the sequence in Figure 5 as the input DNA template. This amplified DNA sequence would contain every triplet combination encoding TTP or modified dUTP insertion, except TTT, and is designed to be challenging for a polymerase. After 30−35 cycles *Pfu* (exo-), D. Vent (exo-), and KOD XL appeared to generate PCR product by real-time analysis. However, upon analysis by denaturing polyacrylamide gel electrophoresis, only short fragments and traces of the full-length DNA were being produced (Figure 5). Disappointingly, none of the PCR enzymes tested were able to generate full-length product using any of the modified dUTP derivatives. In the positive control experiment, *Pfu* (exo-), *Taq*, D. Vent (exo-), or KOD XL with the same template and TTP were all able to generate full-length PCR product between 5 and 10 cycles using real-time analysis, which was confirmed by denaturing polyacrylamide gel electrophoresis analysis. The lack of enzymatic incorporation of these new modified dUTP derivatives was of serious concern, and it was unclear if these new triphosphates were of any use at all.

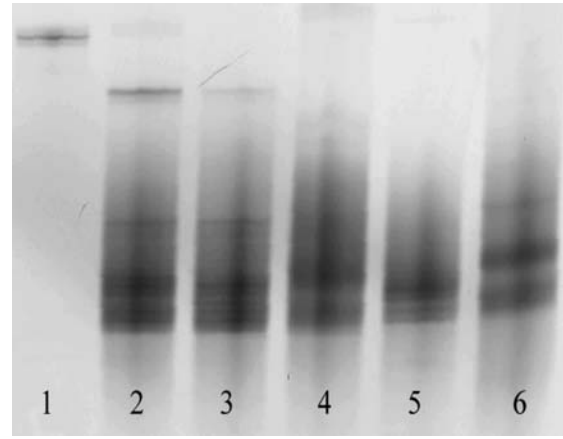


Figure 5. SYBR Gold stained gel image of PCR using fixed ssDNA template (5′-ATA TAT ATG ATG TGA GTG TGT GAC GAG CTC AAT TAT GGT GTT GCC TCT TCC TAG TCG TAC TGA TGC ATC CTC TTC CAC CAC AAC CGA GAC ACA AAA AAA A-3′), KOD XL, dATP, dGTP, dCTP, and modified dUTPs **6a−e** (lanes 2−6, respectively) or TTP (lane 1).

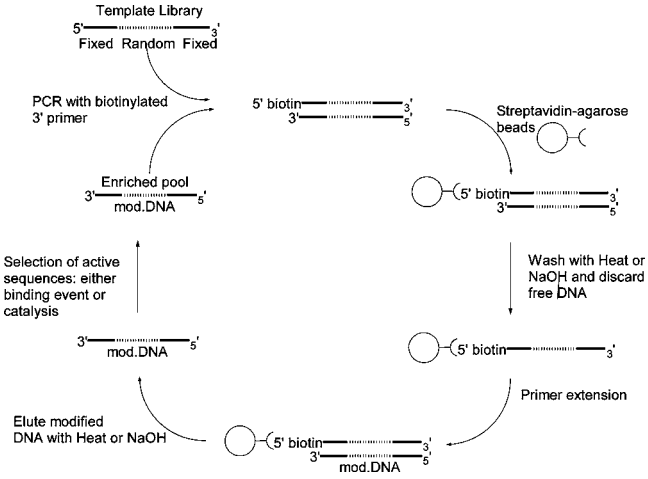
It seemed reasonable to propose that PCR failed because the polymerase must use a modified DNA template for subsequent cycles of PCR amplification after the first extension step, in contrast to the primer extension reactions already discussed above where the polymerases are only required to use an unmodified DNA template and only need to contend with the modified dUTP substrate. It appeared possible that using a modified template was much more difficult for these polymerases than incorporation of these modified dUTP derivatives using an unmodified DNA template.

Primer Extension of Modified Templates Containing 6a−f. In order to test the possibility that primer extension could be successful with TTP instead of modified dUTP to generate unmodified DNA using a modified template, primer extension was performed on six modified DNA templates prepared by KOD primer extension using **6a−f**. Because KOD XL and D. Vent were able to effectively incorporate **6a−f** during primer extension, they were tested for their ability to read a modified 92-mer template containing a 40N random region and synthesize the unmodified complement DNA strand. A biotinylated 3′-primer was used so that the newly synthesized unmodified DNA complement could be treated with streptavidin and separated from the modified template strand on a polyacrylamide gel by electrophoresis. Both KOD XL and D. Vent were able to synthesize the unmodified complement strand. However, it was found that increasing the time of primer extension to 30 min was required to observe excellent yields of full-length DNA from these modified templates. The data obtained from polyacrylamide gel electrophoresis and staining with Sybr Gold are summarized in Table 2. Modified templates containing the napthyl (**6c**), and imidazole (**6f**) groups gave the lowest yield of full-length DNA for both KOD XL and D. Vent polymerase. In addition, the template containing the indole group (**6d**) proved to be a difficult template for KOD giving a modest yield of

Table 2. Yields of Full-Length Products from Primer Extension Reactions on Modified DNA Templates Derived from **6a**−**f** (Yields Represent the Ratio of Streptavidin-Shifted Product to Modified Template As Quantified by Fluorescence Using a Fluorescent Gel Imager)

triphosphate	D. Vent (exo-)	KOD XL
TTP	87%	86%
6a	89%	71%
6b	107%	65%
6c	42%	42%
6d	82%	44%
6e	89%	83%
6f	61%	45%

Scheme 3. Revised DNA Selection Scheme To Allow Use of Modified DNA



full-length DNA. Overall, these data suggest that D. Vent is somewhat better than KOD XL in primer extension to produce DNA from these modified templates.

Revised DNA In Vitro Selection Cycle. Because primer extension was shown to work on these modified templates to give good to excellent yields of full-length DNA, it seemed plausible to construct a revised DNA selection scheme that obviated the need to perform PCR with the modified dUTP derivatives **6a**−**f** (Scheme 3). As in Scheme 2, the selection cycle of Scheme 3 begins with a DNA template library prepared by automated DNA synthesis, and Scheme 3 is identical to Scheme 2 in steps taken until the heat or NaOH wash step to remove the DNA from the beads. At this point in the selection cycle, in contrast to Scheme 2, the free ssDNA would be discarded, leaving the beads with the complementary ssDNA library attached. The beads would then be used for a primer extension step using modified dUTP, dATP, dCTP, and dGTP. This would then generate a modified DNA strand that could subsequently be removed from the beads by another heat or NaOH wash step to isolate the modified ssDNA library. We have demonstrated that the amide linkage of these modified dUTP derivatives is stable to basic conditions for extended periods of time at ambient temperature. This modified ssDNA library could then undergo the selection step for binding or catalysis. The selected sequences would then go into PCR with TTP, dATP, dCTP, and dGTP, avoiding PCR amplification using the modified dUTP. Using the same 3′-primer containing a 5′-biotin would create an enriched library ready to begin the next cycle of selection.

From the results presented above, it was clear that PCR amplification of a modified template with TTP would be

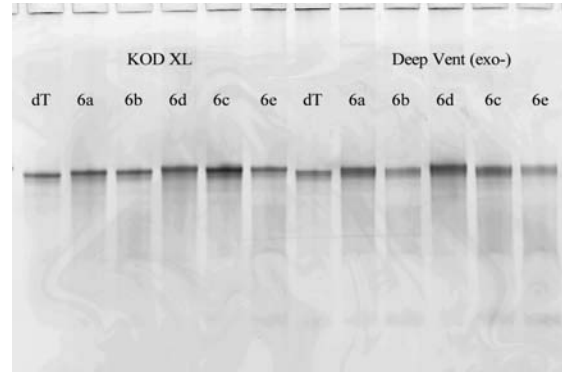


Figure 6. Fluorescent image of Sybr Gold stained primer extension products from a 92 nucleotide ssDNA 40N random template attached to agarose beads using either KOD XL or D. Vent (exo-) polymerase. dT is the TTP control lane. Lanes 6a−6e are primer extension products using modified dUTP derivatives **6a**−**f** shown in Figure 1.

successful if the first PCR step was given a 30 min extension time. Further, as discussed above, primer extension reactions in solution were observed to work well with KOD XL and D. Vent polymerases. However, Scheme 3 requires not only that primer extension occur using modified dUTP derivatives but also that it works using a template strand bound to a bead. Therefore, we prepared a biotinylated template strand (92-mer, 40N random sequence) attached to a streptavidin-coated agarose bead (Pierce), as described in Scheme 3. Primer extension reactions with **6a**−**f** (or TTP), dATP, dCTP, dGTP, and either D. Vent or KOD XL were then performed on the agarose bead immobilized antisense DNA strand. The modified DNA library was then eluted off of the beads using NaOH (20 mM), neutralized with HCl (80 mM), and analyzed by polyacrylamide gel electrophoresis using Sybr Gold staining and fluorescent imaging (Figure 6). Both KOD XL and D. Vent were able to effectively incorporate all of the modified dUTP derivatives **6a**−**f** during these bead-bound primer extension reactions. In accord with the initial primer extension experiments, KOD XL and D. Vent were able to generate full-length products at similar yields to TTP with all six modified dUTP analogues. Taken together, these results presented here set the stage for new DNA in vitro selection experiments utilizing the chemical modifications described herein.

Selection of DNA versus Modified DNA Aptamers. With the in vitro selection methods established as described above, it was of interest to determine if, in fact, modified DNA aptamers to proteins could be selected. To make a challenging test case comparison between DNA and modified DNA, we chose tumor necrosis factor receptor super family member 9 (TNFRSF9), a difficult protein target that in previous DNA selections failed to yield an aptamer with even poor binding (K_d > 100 nM; S. Wilcox, unpublished results). TNFRSF9 is a member of the TNF receptor super family and contributes to the clonal expansion, survival, and development of human T cells. It has a pI of 8.16 and a molecular weight of 26.2 kDa. It is an important therapeutic target for a variety of diseases in cancer and inflammation.[49] Notably, a RNA aptamer for this target has been reported ($K_d = 4 \times 10^{-8}$ M), but DNA aptamers for this important human protein have not.[50] As a positive control for

(49) Croft, M. *Nat. Rev. Immunol.* **2009**, *9*, 271–285.
(50) McNamara II, J. O.; Kolonias, D.; Pastor, F.; Mittler, R. S.; Chen, L.; Giangrande, P. H.; Sullenger, B.; Gilboa, E. *J. Clin. Invest.* **2008**, *118*, 376–386.

Table 3. Measured K_d Values for Selected DNA Pools Containing Either TTP, **6a**, **6b**, or **6d**

target protein	TTP (K_d, M)	**6a** (K_d, M)	**6b** (K_d, M)	**6d** (K_d, M)
TNFRSF9	$>1 \times 10^{-7}$	6×10^{-9}	$>1 \times 10^{-7}$	4×10^{-9}
TACSTD2	9×10^{-9}	3×10^{-9}	5×10^{-9}	5×10^{-10}

the DNA in vitro selection methods described above, tumor-associated calcium signal transducer 2 (TACSTD2) was also used as a protein target in a parallel experiment. TACSTD2 is known to bind random DNA pools with a $K_d < 10^{-7}$ M and appeared to be a good target for selection of DNA aptamers, although it should be noted that no DNA aptamers had been reported for this target. The starting random DNA libraries were prepared using standard automated phosphoramidite synthesis on an ABI 3900 synthesizer to give a base composition ratio of ca. 1:1:1:1 A:C:G:T in the 40N random region. These aptamer selection experiments on the protein targets TNFRSF9 and TACSTD2 were monitored by C_0t curve analysis for convergence of the sequence pools. Protein concentrations for the first cycle of selection were on the order of 1 μM and DNA library concentration 1 nM. Selection for affinity binding was driven by decreasing protein concentration during the course of the selection. Both selection experiments were stopped at 8 cycles based on convergence.

The PCR product of the cycle 8 pool was converted into the ssDNA aptamer pool by primer extension methods using KOD as described above. Analysis of these experiments began by measuring the K_d of the DNA and modified DNA aptamer pools using radiolabeled pools in a filter binding assay. A cutoff of 100 nM was used for the K_d measurement with a measured K_d above this value being considered a failure. In support of the notion that TACSTD2 was a suitable control protein, all of the DNA selections (TTP, **6a**, **6b**, and **6d**) yielded aptamers. Good ($K_d = 9$ nM) to excellent ($K_d = 0.5$ nM) binding affinity was observed for this target protein (Table 3). As had been observed previously, the DNA library derived from TTP resulted in a failed selection for TNFRSF9. The selection using the DNA library derived from **6b** also failed on this target. Gratifyingly, selections performed with DNA libraries derived from **6a** and **6d** yielded sequence pools with K_d values well below the cutoff limit of 100 nM (10^{-7} M). Perhaps most surprisingly, the DNA aptamer pool derived from **6d** had a significantly lower K_d for TNFRSF9 than that reported for the unmodified RNA aptamer (Table 3 and Figure 7). These data clearly show the advantage of using DNA base modifications at the 5-postion of the uridine ring in aptamers.

However, it could not be ruled out from the pool data that the types of modifications used for the aptamers described herein were simply hydrophobic and that the specific structure of the 5-position modification was less important. Moreover, without sequencing and performing base composition analysis on the clones, it was unclear if the modified dUTP analogues were even tolerated in the structures or of any importance whatsoever. To test this possibility, a representative "hit" sequence was examined for binding to the target protein TNFRSF9.

Cloning and sequencing of the cycle 8 aptamer pool from the **6d** selection on the TNFRSF9 target was performed and revealed two distinct families. From a sequencing perspective, this selection resulted in a high degree of convergence, making analysis of the major family descriptive of the outcome of the in vitro selection. DNA clone 1684-40 was chosen for further study because it was a member of a major sequence family. Its

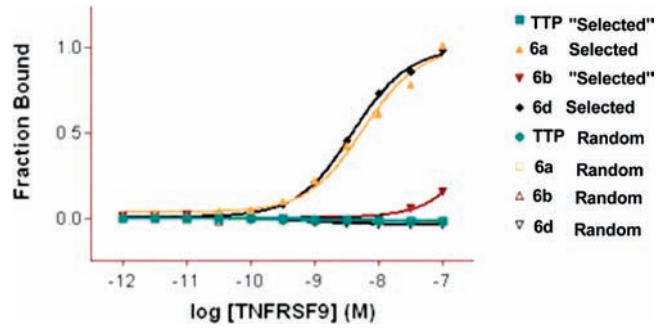


Figure 7. Plot of filter binding data for K_d measurements of the in vitro selected pools and the starting random libraries using either TTP, **6a**, **6b**, or **6d** for TNFRSF9. Note that for TTP "selected" and TTP random, no binding was observed in either case. Fraction bound was calculated as the fraction of radiolabeled DNA bound to protein immobilized on beads as a function of protein concentration and were normalized to the amount of DNA bound at the curve plateau.

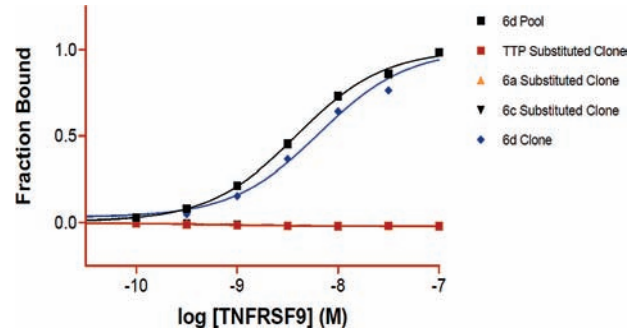


Figure 8. Plot of filter binding data for K_d measurements of the in vitro selected DNA sequence clone 1684-40 and the evolved cycle 8 pool for the protein target TNFRSF9 using either TTP, **6a**, or **6d** for enzymatic preparation. Fraction bound was calculated as the fraction of radiolabeled DNA bound to protein immobilized on beads as a function of protein concentration and normalized to the amount of DNA bound at the curve plateau.

sequence was C C T G C A C C C A G T G T C C C ***C G A C G G G G C C C T C T A G C C G T A C T C T G T A A T G G C G G A T G C T G A C*** G G A G A G G A G G A C G G, where the letters in bold italics indicate the sequence evolved from the random region. The base composition of this sequence was representative of the aptamers in this family at 31% C, 21% T (modified dU), 31% G, and 17% A. Note that the sequence was prepared enzymatically by primer extension using KOD XL as described above, which means that only positions labeled italicized *T* will contain a modification within the DNA sequence.

To determine if the binding to TNFRSF9 was merely a consequence of adding some hydrophobic groups to clone 1684-40 or a more specific structural effect unique to the modification used in the selection of this sequence, primer extension reactions were performed using either **6d** (indole dUTP analogue used in the selection), **6a** (benzyl), or **6c** (napthyl) dUTP analogues. In addition, clone 1684-40 was also transcribed with TTP to determine if this sequence required modification at all to bind TNFRSF9. Employing the same filter binding assay as used to study the evolved DNA aptamer pool binding, and including a control experiment with the cycle 8 aptamer pool, the K_d values were compared and contrasted, as shown in Figure 8. No K_d was observed within the 100 nM cutoff limits for the DNA aptamer sequence 1684-40 binding TNFRSF9 when prepared from TTP (red squares). Similarly, clone 1684-40 prepared by

enzymatic transcription with either **6a** or **6c** gave no binding to TNFRSF9 within the 100 nM cutoff. In contrast, enzymatic preparation of 1684-40 with **6d** gave a K_d of approximately 5 nM in binding TNFRSF9, within experimental error, the same as that observed for the evolved cycle 8 pool. It is somewhat surprising that substitution of **6c** (napthyl) for another bicyclic aromatic **6d** (indole) did not give an aptamer with at least weak binding to TNFRSF9. Perhaps these aromatic groups confer different aptamer structures, or these apparent subtle differences in pendant aromatic groups confer more significant changes at the aptamer protein interface, or both. Of course, it is true that indole is not just another aromatic and can form another H-bond to a target, whereas napthyl cannot. It may be that the indole groups in clone 1684-40 form important contacts with TNFRSF9.

Further spectroscopic work is underway to determine the molecular basis for the apparent binding improvements conferred by these modified DNA aptamers for a refractory protein target TNFRSF9, as well as the control protein target TACSTD2. In addition, we hope to learn if aptamer structures may be altered by the presence of these modifications. In any event, we are encouraged by these successful in vitro selections against refractory protein targets, and significant effort is now underway to select aptamers for a wide range of targets in the human proteome.

Conclusions

Streamlined synthetic methods are now available to prepare 5-position modified dUTP derivatives analogous to the UTP derivatives previously used to successfully identify new RNA catalysts for organic reactions via in vitro selection. PCR amplification typically used for DNA in vitro selection was unsuccessful. Fortunately, the dUTP derivatives of this study were incorporated efficiently in a primer extension step, and a modified in vitro selection method was developed that enables the use of these dUTP derivatives for in vitro selection. Of the DNA polymerases studied, the Family B polymerases, D. Vent and KOD XL, were most successful in generating full-length DNA. The modified DNA template could also be used in primer extension using TTP to create unmodified DNA suitable for standard PCR using these polymerases. The ability to perform primer extension with **6a**−**f** on streptavidin agarose beads demonstrates the final step for a revised DNA in vitro selection that avoids performing PCR with modified dUTP derivatives.

With these new DNA modification chemistries established, it will be possible to perform in vitro selection experiments using DNA that were previously successful with RNA. There are many apparent (2′-hydroxyl) and less apparent (pseudorotaion of ribose) differences between RNA and DNA structure and dynamics. It is unknown if these structural differences make RNA a superior catalyst or aptamer when compared to DNA. However, for the protein target TNFRSF9, the modified DNA aptamer described herein had a lower K_d when compared to that of the published RNA aptamer (approximately 5 versus 48 nM). Having these chemistries and in vitro selection methods defined expands the types of modifications available for the DNA aptamer field and paves the way to compare these two related nucleic acid biopolymer platforms.

The DNA modification chemistry and associated in vitro selection methods described herein are now being applied in a large-scale effort to select aptamers to the estimated 3500 proteins at the core of the circulating human proteome to use in multiplexed arrays for human proteome analysis. The results of this effort will show the extent to which these new DNA aptamers can serve as useful reagents in this endeavor, the ultimate goal of which is to discover proteomic biomarkers to help detect, diagnose, and treat human disease. It is anticipated that modified DNA aptamers will be generated by the chemistries and methods described herein to greater than 1000 human protein targets by the end of 2010.

Experimental Section

General. All compounds were prepared from reagent grade starting materials as purchased from Aldrich unless otherwise noted. Tetrakis[triphenylphosphine]palladium(0) was used as received from Strem Chemicals. NMR spectra were recorded on Varian 300 or 400 MHz spectrometers and are reported as δ values, referenced to solvent resonances. ESI mass spectra were recorded at SomaLogic (Boulder, CO).

5′-Primer (5′-ATATATATGATGTGAGTGTGTGACGAG-3′) was obtained from IDT and used without further purification. Biotinylated primer (5′-B-TTTTTTTTGTGTCTCGGTTGTGGTG-3′), Cy-5-labeled primer (5′-Cy5-TTTTTTTTGTGTCTCGGTTGTGGTG-3′), 40N random template (5′-B-TTTTTTTTGTGTCTCGGTTGTGGTG40NCTCGTCACACACTCACATCATATATAT), and triplet dU template (TTTTTTTTGTGTCTCGGTTGTGGTG-GAA GAG GAT GCA TCA GTA CGA CTA GGA AGA GGC AAC ACC ATA ATT GAG-CTCGTCACACACTCACATCATATATAT), where B and Cy-5 represent either a biotin group or Cy-5 group appended during DNA synthesis using biotin phosphoramidite or Cy-5 phosphoramidite from Glen research, were prepared using standard phosphoramidite chemistry on an ABI 394, reversed-phase HPLC purified, then gel purified. Streptavidin used for gel shift analysis of primer extension products was dissolved in phosphate buffered saline. UltraLink Immobilized Streptavidin Plus agarose beads (Pierce) were washed 3 times with 1× wash buffer (100 mM NaCl, 40 mM HEPES pH 7.5, 5 mM KCl, 1 mM $MgCl_2$, and 0.25% Tween-20) and resuspended to their original volume as a 50% slurry immediately before use. All gels of unlabeled DNA were visualized and quantitated by staining with SYBR Gold nucleic acid gel stain (Invitrogen) and imaged on a FujiFilm 5100 fluorescent imager.

Synthesis of Nucleoside Precursors, 5a−5f. In a heavy-walled glass vessel (200 mL) equipped with a Teflon vacuum stopcock, 5-iodo-5′-O-DMTr-3′-O-acetyl-2′-deoxyuridine (0.72 mmol), 3 equiv of amine, 5 equiv of triethylamine, 0.1 equiv of tetrakis-[triphenylphosphine]palladium(0), and 15 mL of 1:1 DMA/THF (freshly distilled off of sodium/potassium alloy and benzophenone) were added. The vessel was evacuated and filled with 50 psi of CO three times, then sealed and heated at 70 °C for 48−72 h. The reaction was cooled, vented, and evaporated under high vacuum to a bubbly yellow solid. The resulting solid was purified with flash silica gel using 1:20 MeOH/CH_2Cl_2. The nucleoside (ca. 95% pure by ^{1}H NMR) was then dissolved a 1:1 solution of MeOH/CH_2Cl_2; 1.1 equiv of 3% TCA (in CH_2Cl_2) was added, and the solution was stirred at room temperature for 2−4 h. The reaction mixture was evaporated to dryness under high vacuum, and the resulting bright orange residue was purified with flash silica gel using 5−10% MeOH in CH_2Cl_2. Products were then used as recovered or recrystallized from ethyl acetate, yielding white to slightly yellowish crystals.

Spectroscopic Data for Nucleosides 5a−f.

5a: ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.90 (s, 1H), 9.11 (s, 1H), 8.78 (s, 1H), 7.27 (m, 5H), 6.14 (t, J = 7 Hz, 1H), 5.22 (s, 1H), 5.19 (s, 1H), 4.47 (d, J = 6 Hz, 2H), 4.09 (s, 1H), δ 3.61 (s, 2H), 2.34 (m, 2H), 2.05 (s, 3H); ^{13}C NMR (300 MHz, DMSO-d_6) δ 170.7, 164.0, 163.8, 150.7, 146.7, 141.6, 138.1, 128.9, 127.5, 127.2, 105.9, 87.0, 86.1, 75.4, 61.8, 38.3, 21.5.

5b: ^{1}H NMR (300 MHz, DMSO-d_6) δ 8.75 (s, 1H), 8.71 (s, 1H), 6.11 (t, J = 7 Hz, 1H), 5.20 (s, 1H), 5.15 (s, 1H), 4.06 (s, 1H), 3.58 (s, 2H), 3.30 (s, 1H), 3.07 (t, J = 6 Hz, 2H), 2.32 (m, 2H),

2.02 (s, 3H), 1.73 (m, 1H), 0.84 (d, $J = 6$ Hz, 6 H); ^{13}C NMR (300 MHz, DMSO-d_6) δ 170.7, 164.0, 162.0, 150.2, 146.5, 106.3, 86.1, 86.0, 75.4, 61.8, 46.4, 38.3, 28.8, 21.5, 20.7.

5c: ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.92 (s, 1H), 9.14 (t, $J = 6$ Hz, 1H), 8.80 (s, 1H), 8.09 (d, $J = 8$ Hz, 1H), 7.94 (d, $J = 8$ Hz, 1H), 7.85 (m, 1H), 7.74 (m, 1H), 7.53 (m, 3H), 6.13 (t, $J = 7$ Hz, 1H), 5.22 (m, 1H), 4.93 (d, $J = 6$ Hz, 2H), 4.10 (d, $J = 4$ Hz, 1H), 3.61 (s, 2H), 2.35 (m, 2H), 2.05 (s, 3H); ^{13}C NMR (400 MHz, DMSO-d_6) δ 170.7, 163.9, 162.0, 150.2, 146.8, 135.7, 134.0, 131.5, 129.3, 128.2, 127.1, 126.6, 126.3, 126.2, 124.0, 121.2 106.1, 86.2, 86.0, 75.4, 61.9, 38.3, 21.5.

5d: ^{1}H NMR (300 MHz, DMSO-d_6) δ 11.87 (s, 1H), 10.81 (s, 1H), 8.80 (t, $J = 8$ Hz, 1H), 8.74 (s, 1H), 7.56 (d, $J = 8$ Hz, 1H), 7.32 (d, $J = 8$ Hz, 1H), 7.14 (d, $J = 2$ Hz, 1H), 7.05 (t, $J = 7$ Hz, 1H), 6.95 (t, $J = 8$ Hz, 1H), 6.13 (t, $J = 7$ Hz, 1H), 5.22 (s, 1H), 5.18 (t, $J = 8$ Hz, 1H), 4.07 (s, 1H), 3.61 (s, 2H), 3.54 (dd, $J = 13$, 7 Hz, 2H), 2.89 (t, $J = 7$ Hz, 2H), 2.31 (m, 2H), 2.05 (s, 3H); ^{13}C NMR (300 MHz, DMSO-d_6) δ 170.7, 163.7, 162.0, 150.3, 146.7, 141.4, 136.9, 127.8, 123.5, 123.3, 121.6, 119.0, 112.2, 112.1, 106.4, 86.1, 86.0, 75.4, 61.9, 38.3, 25.9, 21.5.

5e: ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.90 (s, 1H), 9.09 (t, $J = 6$ Hz, 1H), 8.45 (m, 2H), 8.41 (s, 1H), 7.57 (m, 2H), 6.10 (t, $J = 7$ Hz, 1H), 5.20 (s, 1H), 5.17 (s, 1H), 4.44 (d, $J = 6$ Hz, 2H), 4.09 (d, $J = 4$ Hz, 1H), 3.51 (s, 2H), 2.31 (m, 2H), 2.05 (s, 3H); ^{13}C NMR (300 MHz, DMSO-d_6) δ 170.7, 164.0, 163.8, 150.7, 146.7, 141.6, 138.1, 128.9, 127.5, 127.2, 105.9, 87.0, 86.1, 75.4, 61.8, 38.3, 21.5.

5f: ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.85 (s, 1H), 9.44 (s, 1H), 8.53 (s, 1H), 8.30 (s, 1H), 7.49 (s, 1H), 6.77 (s, 1H), 6.15 (t, $J = 8$ Hz, 1H), 5.21 (d, $J = 3$ Hz, 2H), 4.01 (s, 1H), 3.59 (ddd, $J = 16$, 12, 4 Hz, 2H), 3.44 (dd, $J = 13$, 7 Hz, 2H), 2.66 (t, $J = 7$ Hz, 2H), 2.25 (m, 2H), 2.04 (s, 3H); ^{13}C NMR (400 MHz, DMSO-d_6) δ 170.7, 164.5, 163.8, 154.0, 144.9, 135.3, 129.4, 106.3, 94.6, 85.8, 85.4, 79.9, 75.6, 62.0, 38.1, 29.9, 21.6.

Synthesis of UTP Derivatives 6a−f. All 5-position modified dUTP derivatives were prepared as described by the method of Eckstein.[36]

Spectroscopic Data for dUTPs 6a−f.

6a: UV λ_{max} 280 nm ε 13 700 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.38 (s, 1H), 7.19 (m, 5H), 6.06 (t, $J = 7$ Hz, 1H), 4.46 (m, 1H), 4.39 (d, $J = 6$ Hz, 2H), 4.08 (dd, $J = 7$, 4 Hz, 1H), 4.02 (m, 2H), 2.28 (dd, $J = 10$, 4 Hz, 2H); ^{31}P NMR (400 MHz, D_2O) δ −9.64 (d, $J = 51$ Hz, 1P), 10.40 (d, $J = 52$ Hz, 1P), 22.28 (t, $J = 48$ Hz, 1P); HRMS (ESI) calcd for $C_{17}H_{21}N_3O_{15}P_3$ (M$^-$) 600.0191, found 600.0197.

6b: UV λ_{max} 276 nm ε 10 200 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.41 (s, 1H), 6.09 (t, $J = 7$ Hz, 1H), 4.47 (m, 1H), 4.10 (m, 1H), 4.04 (m, 2H), 2.27 (m, 2H), 1.71 (m, 2H), 0.75 (d, $J = 6$ Hz, 6H); ^{31}P NMR (400 MHz, D_2O) δ −9.64 (d, $J = 51$ Hz, 1P), −10.40 (d, $J = 52$ Hz, 1P), −22.28 (t, $J = 48$ Hz, 1P); HRMS (ESI) calcd for $C_{14}H_{23}N_3O_{15}P_3$ (M$^-$) 566.0347, found 566.0354.

6c: UV λ_{max} 281 nm ε 20 000 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.22 (s, 1H), 7.79 (d, $J = 8$ Hz, 1H), 7.74 (d, $J = 7$ Hz, 1H), 7.66 (d, $J = 8$ Hz, 1H), 7.39 (m, 2H), 7.29 (m, 2H), 5.93 (t, $J = 7$ Hz, 1H), 4.70 (dd, $J = 27$, 16 Hz, 2H), 4.42 (m, 1H), 4.08 (m, 1H), 4.02 (m, 2H), 2.26 (m, 1H), 2.18 (m, 1H); ^{31}P NMR (400 MHz, D_2O) δ −10.22 (d, $J = 46$ Hz, 1P), −10.62 (d, $J = 50$ Hz, 1P), −22.67 (t, $J = 44$ Hz, 1P); HRMS (ESI) calcd for $C_{21}H_{23}N_3O_{15}P_3$ (M$^-$) 650.0347, found 650.0365.

6d: UV λ_{max} 279 nm ε 13 100 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.22 (s, 1H), 7.79 (d, $J = 8$ Hz, 1H), 7.74 (d, $J = 7$ Hz, 1H), 7.66 (d, $J = 8$ Hz, 1H), 7.39 (m, 2H), 7.29 (m, 2H), 5.93 (t, $J = 7$ Hz, 1H), 4.70 (dd, $J = 27$, 16 Hz, 2H), 4.42 (m, 1H), 4.08 (m, 1H), 4.02 (m, 2H), 2.26 (m, 1H), 2.18 (m, 1H); ^{31}P NMR (400 MHz, D_2O) δ −10.22 (d, $J = 46$ Hz, 1P), −10.62 (d, $J = 50$ Hz, 1P), −22.67 (t, $J = 44$ Hz, 1P); HRMS (ESI) calcd for $C_{20}H_{24}N_4O_{15}P_3$ (M$^-$) 653.0456, found 653.0450.

6e: UV λ_{max} 277 nm ε 12 100 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.44 (m, 2H), 8.40 (s, 1H), 7.57 (m, 2H), 6.04 (t, $J = 7$ Hz, 1H), 4.64 (s, 1H), 4.58 (s, 2H), 4.44 (m, 1H), 4.07 (d, $J = 4$ Hz, 2H), 3.99 (m, 2H), 2.29 (m, 2H); ^{31}P NMR (400 MHz, D_2O) δ −9.80 (d, $J = 46$ Hz, 1P), −10.40 (d, $J = 50$ Hz, 1P), −22.33 (t, $J = 44$ Hz, 1P); HRMS (ESI) calcd for $C_{16}H_{20}N_4O_{15}P_3$ (M$^-$) 601.0143, found 601.0169.

6f: UV λ_{max} 276 nm ε 13 500 cm^{-1} M^{-1}; ^{1}H NMR (400 MHz, D_2O) δ 8.36 (s, 1H), 8.27 (s, 1H), 7.04 (s, 1H), 6.05 (t, $J = 7$ Hz, 1H), 4.47 (m, 1H), 4.09 (d, $J = 2$ Hz, 1H), 4.02 (dd, $J = 10$, 7 Hz 2H), 3.49 (ddd, $J = 9$, 6, 3 Hz, 2H), 2.82 (t, $J = 6$ Hz, 2H), 2.28 (m, 2H); ^{31}P NMR (400 MHz, D_2O) δ −8.14 (d, $J = 48$ Hz, 1P), −10.32 (d, $J = 47$ Hz, 1P), −21.82 (t, $J = 49$ Hz, 1P); LRMS (ESI) calcd for $C_{15}H_{21}N_5O_{15}P_3$ (M$^-$) 604.3, found 604.3

Primer Extension Reactions for Enzyme Screening with 6a−f. *Pfu* (exo-) and KF (exo-) were purchased from Stratagene; *Taq* was purchased from Abgene; KOD XL was purchased from Novagen; Deep Vent (exo-) was purchased from New England Biolabs, and *Tth* was purchased from Epicenter Biotechnologies. Reaction buffers were used as received by the supplier of the DNA polymerase except for KF, which was prepared according to the test conditions described in the product insert and consisted of 66 mM $KHPO_4$ (pH 7.4), 10 mM $MgCl_2$, and 1 mM β-mercaptoethanol. 5′-Primer either labeled with biotin or with Cy-5 (final concentration 3.75 μM) and 40N template (final concentration 0.75 μM) was mixed in 1× reaction buffer and heated to 95 °C for 1 min and then cooled to room temperature over 10 min. DNA polymerase (1.5 U), dATP, dCTP, dGTP, TTP or one of the modified dUTP derivatives **6a−f** (final concentration 200 μM each dNTP), and $MgCl_2$ (to *Pfu*, *Tth*, and *Taq* as specified in each product insert) were added to a final volume of 15 μL. Reactions were carried out at 70 °C for 30 min, except for KF reactions, which were carried out 37 °C. For reactions analyzed by gel shift, streptavidin was immediately added to each reaction at 10-fold molar excess to template. After 1 min, 15 μL of denaturing gel-loading buffer (8 M urea, 20 mM Tris pH 8.0, 1 mM EDTA, and 0.05% bromophenol blue) was added to each reaction and heated to 70 °C for 2 min. Reactions were analyzed by 8% PAGE (19:1 acrylamide/bisacrylamide, 8 M urea, 45−50 °C), visualized, and quantified by fluorescence.

Primer Extension Reactions with 6a−f on Streptavidin Agarose Beads. A 40N random template or triplet dU template was amplified by PCR with 3′-biotinylated primer and 5′-primer. The resulting dsDNA was immobilized on streptavidin agarose beads (Pierce) by adding 15 μL of a 50% bead slurry per 50 μL PCR reaction. Reactions were combined and shaken at room temperature for 5 min before being spun in a centrifuge for 20 s. Supernatant was removed, and then the beads were washed three times with 1× wash buffer by shaking, spinning, and removing supernatant. ssDNA was then eluted from the beads using 80 mM NaOH and neutralized with 20 mM HCl. The beads were washed three times with 1× wash buffer and resuspended in 1× primer extension buffer consisting of 1× reaction buffer (as supplied by polymerase manufacturer), 1.5 U KOD XL or D. Vent, 5′-primer (final concentration 4 μM), dATP, dCTP, dGTP, and TTP or one of the modified dUTP derivatives **6a−f** (final concentration 400 μM each dNTP) at a final volume of 50 μL. Two 50 μL PCR reactions (30 μL of orginal 50% bead slurry) were used per single 50 μL primer extension reaction. Reactions were then heated to 95 °C for 30 s, annealed at 55 °C for 15 s, heated at 70 °C for 30 min with shaking, and then spun in a centrifuge for 20 s. Supernatant was removed and each reaction washed three times with 1× wash buffer as above. Modified ssDNA was then eluted from the beads using 80 mM NaOH and neutralized with 20 mM HCl. Beads were then washed three times with 1× wash buffer before being used again. Beads were used four times for primer extension reactions with no detectable decrease in product yield. Eluted products were then diluted in 2× denaturing gel-loading buffer (8 M urea, 20 mM Tris pH 8.0, 1 mM EDTA, and 0.05% bromophenol blue), analyzed by 8% denaturing PAGE, and visualized by fluorescence.

Primer Extension Reactions with dATP, dCTP, dGTP, and TTP on Templates Modified with 6a−f. Modified 40N template was enzymatically synthesized on beads and purified as described above, mixed with 3′-biotinylated primer in 1× reaction buffer (as supplied by polymerase manufacturer), heated to 95 °C for 1 min, cooled to room temperature over 10 min, and then subjected to reaction conditions identical to those described above for the primer extension reactions used for the enzyme screening, using only KOD XL and D. Vent (exo-).

Real-Time PCR with 6a−f. PCR was performed on a Biorad MyIQ real-time PCR thermal cycler with single-color detection. Reactions were carried out in 50 μL 1× reaction buffer, $MgCl_2$ (an additional 2 mM to that contained in the supplied reaction buffer or an additional 2 mM to the minimal amount specified for the reaction buffers without $MgCl_2$), 1.25 U polymerase, dATP, dCTP, dGTP, and TTP or one of the modified dUTP derivatives **6a−f** (final concentration 200 μM each dNTP), 5′-primer and 3′-biotinylated primer (final concentration 2 μM each), 0.05 pmol triplet dU template, and SYBR Green I (Invitrogen). Thermal cycling conditions consisted of an initial denaturing step at 95 °C for 1 min, 55 °C for 15 s, 70 °C for 1 min, then repeated cycling through these steps only cutting back the denaturing step at 95 °C to 30 s. Products were then diluted in 2× denaturing gel-loading buffer (8 M urea, 20 mM Tris pH 8.0, 1 mM EDTA, and 0.05% bromophenol blue), analyzed by 8% denaturing PAGE, and visualized by fluorescence.

In Vitro Selection. DNA and modified DNA libraries were prepared with dATP, dGTP, 5-methyl-dCTP (MedCTP), and either dTTP or one of three dUTP analogues: **6a**, **6b**, or **6d**. Candidate mixtures were prepared by polymerase extension of a primer annealed to a biotinylated template (5′- ABABCCGTCCTC-CTCTCCGT-40N-GGGACACTGGGTGCAGG-3′), where B indicates a biotin incorporated during DNA synthesis and 40N indicates a 40 nucleotide random region. For each candidate mixture composition, 4.8 nmol forward PCR primer (5′-ATATATATCCT-GCACCCAGTGTCCC-3′) and 4 nmol template were combined in 100 μL 1× KOD DNA polymerase buffer (Novagen), heated to 95 °C for 8 min, and cooled on ice. Each 100 μL primer/template mixture was added to a 400 μL extension reaction containing 1× KOD DNA polymerase buffer, 0.125 U/μL KOD DNA polymerase (Novagen), and 0.5 mM each dATP, MedCTP, dGTP, and dTTP or dUTP analogue and incubated at 70 °C for 30 min. Double-stranded product was captured via the template strand biotins by adding 1 mL of streptavidin-coated magnetic beads (Magna-Bind Streptavidin, Pierce, 5 mg/mL in 1 M NaCl + 0.05% Tween-20) and incubating at 25 °C for 10 min with mixing. Beads were washed three times with 0.75 mL of SB1T buffer (40 mM HEPES, pH 7.5, 125 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.05% Tween-20). The aptamer strand was eluted from the beads with 1.2 mL of 20 mM NaOH, neutralized with 0.3 mL of 80 mM HCl, and buffered with 15 μL of 1 M HEPES, pH 7.5. Candidate mixtures were concentrated with a Centricon-30 to approximately 0.2 mL and quantified by UV absorbance spectroscopy.

The extracellular domains of TNFRSF9 and TACSTD2 fused to IgG1 and 6-His were purchased from R&D Systems. TNFRSF9 and TACSTD2 were immobilized on Co^{2+}-NTA paramagnetic beads (TALON, Dynal). Target proteins were diluted to 0.2 mg/mL in 0.5 mL of B/W buffer (50 mM Na-phosphate, pH 8.0, 300 mM NaCl, 0.01% Tween-20) and added to 0.5 mL of TALON beads (prewashed three times with B/W buffer and resuspended to 10 mg/mL in B/W buffer). The mixture was rotated for 30 min at 25 °C and stored at 4 °C until use. TALON beads coated with $(His)_6$ peptide were also prepared and stored as above. Prior to use, beads were washed three times with B/W buffer, once with SB1T, and resuspended in SB1T.

Affinity selections were performed separately with each candidate mixture, comparing binding between target protein beads (signal) and $(His)_6$ beads (background). For each sample, a 0.5 μM candidate DNA mixture was prepared in 40 μL of SB1T. Ten microliters of a protein competitor mixture was added to each DNA mix (0.1% HSA, 10 μM casein, and 10 μM prothrombin in SB1T).

Binding reactions were performed by adding 50 μL target protein-coated beads or $(His)_6$-coated beads (5 mg/mL in SB1T) to the DNA mixture and incubating at 37 °C for 15 min with mixing. The DNA solution was removed, and the beads were washed five times at 37 °C with SB1T containing 0.1 mg/mL herring sperm DNA (Sigma-Aldrich). Unless indicated, all washes were performed by resuspending the beads in 100 μL wash solution, mixing for 30 s, separating the beads with a magnet, and removing the wash solution. Bound aptamers were eluted from the beads by adding 100 μL of SB1T + 2 M guanidine-HCl and incubating at 37 °C for 5 min with mixing. The aptamer eluate was transferred to a new tube after magnetic separation. After the first two selection rounds, the final two of five target bead washes were done for 5 min instead of 30 s.

Primer beads were prepared by immobilizing biotinylated reverse PCR primer (5′-ABABTTTTTTTTCCGTCCTCCTCTCCGTC-3′) to streptavidin-coated paramagnetic beads (MyOne-SA, Dynal). Five milliliter MyOne-SA beads (10 mg/mL) were washed once with NaClT (5 M NaCl, 0.01% Tween-20) and resuspended in 5 mL of biotinylated reverse PCR primer (5 μM in NaClT). The sample was incubated at 25 °C for 15 min, washed twice with 5 mL of NaClT, resuspended in 12.5 mL of NaClT (4 mg/mL), and stored at 4 °C.

Twenty-five microliter primer beads (4 mg/mL in NaClT) were added to the 100 μL aptamer solution in guanidine buffer and incubated at 50 °C for 15 min with mixing. The aptamer solution was removed, and the beads were washed five times with SB1T. Aptamer was eluted from the beads by adding 85 μL of 20 mM NaOH and incubating at 37 °C for 1 min with mixing. Eighty microliters of aptamer eluate was transferred to a new tube after magnetic separation, neutralized with 20 μL of 80 mM HCl, and buffered with 1 μL of 0.5 M Tris-HCl, pH 7.5.

Selected aptamer DNA was amplified and quantified by QPCR. Forty-eight microliters of DNA was added to 12 μL of QPCR Mix (5X KOD DNA polymerase buffer, 25 mM $MgCl_2$, 10 μM forward PCR primer, 10 μM biotinylated reverse PCR primer, 5× SYBR Green I, 0.125 U/μL KOD DNA polymerase, and 1 mM each dATP, dCTP, dGTP, and dTTP) and thermal cycled in an ABI5700 QPCR instrument with the following protocol: 1 cycle of 99.9 °C, 15 s, 55 °C, 10 s, 70 °C, 30 min; 30 cycles of 99.9 °C, 15 s, 72 °C, 1 min. Quantification was done with the instrument software, and the number of copies of DNA selected with target beads and $(His)_6$ beads was compared to determine signal/background ratios.

Following amplification, the PCR product was captured on MyOne-SA beads via the biotinylated antisense strand; 1.25 mL MyOne-SA beads (10 mg/mL) were washed twice with 0.5 mL of 20 mM NaOH, once with 0.5 mL of SB1T, resuspended in 2.5 mL of 3 M NaCl, and stored at 4 °C. Twenty-five microliter MyOne-SA beads (4 mg/mL in 3 M NaCl) were added to 50 μL double-stranded QPCR product and incubated at 25 °C for 5 min with mixing. The beads were washed once with SB1T, and the "sense" strand was eluted from the beads by adding 200 μL of 20 mM NaOH and incubating at 37 °C for 1 min with mixing. The eluted strand was discarded, and the beads were washed three times with SB1T and once with 16 mM NaCl. Aptamer sense strand was then prepared with the appropriate nucleotide composition as described above and the whole cycle repeated.

***K_d* Measurements.** Affinities of the enriched libraries were measured using TALON bead partitioning. DNA was renatured by heating to 95 °C and slowly cooling to 37 °C. Complexes were formed by mixing a low concentration of radiolabeled DNA (~1×10^{-11} M) with a range of concentrations of target protein (1×10^{-7} to 1×10^{-12} M final) in SB1 buffer and incubating at 37 °C.

A portion of each reaction was transferred to a nylon membrane and dried to determine total counts in each reaction. Twenty-five micrograms of MyOne TALON beads (Invitrogen) was added to the remainder of each reaction and mixed at 37 °C for 1 min. A portion was then passed through a MultiScreen HV plate under vacuum to separate protein-bound complexes from unbound DNA and washed with 100 μL of SB1 buffer. The nylon membranes and MultiScreen HV plates were phosphorimaged, and the amount of radioactivity in each sample was quantified using a FUJI FLA-3000. The fraction of captured DNA was plotted as a function of protein concentration, and a nonlinear curve-fitting algorithm was used to extract equilibrium binding constants (K_d values) from the data.

Acknowledgment. We thank the U.S. Department of Energy (DOE), Office of Basic Energy Science (BES), and the National Science Foundation (CHE 0414527) for partial financial support of this work.

JA908035G