J|A|C|S
ARTICLES
Published on Web 07/21/2006

A Binary DNA Probe for Highly Specific Nucleic Acid Recognition

Dmitry M. Kolpashchikov

Contribution from the Division of Experimental Therapeutics, Department of Medicine, Columbia University, Box 84, 630W 168th Street, New York, New York 10032

Received April 21, 2006; E-mail: dk2111@clumbia.edu

Abstract: A new concept for nucleic acid probe design is reported. The extremely high selectivity of the probe is predetermined by cooperative hybridization of the two relatively short (10 nucleotide) DNA hairpin fragments to the analyte. A binary DNA probe fluorescently reports the presence of 0.5% of the analyte in excess amount of a single base substituted oligodeoxyribonucleotide and distinguishes single nucleotide substitutions at any position of a 20-mer oligonucleotide at room temperature.

Numerous techniques for DNA/RNA analysis rely on the ability of the probe to recognize nucleic acid sequences specifically by forming duplexes. The formation of at least 15−20 nucleotide hybrids between the probe and the analyte is required to uniquely define a specific fragment in a nucleic acid the size of a genome. Hybrids of such length are too stable to be sensitive to a base mispairing, since a single mismatch unit results in a relatively small energetic penalty.[1] Several strategies have been developed to solve this problem. A general approach employs destabilization of the duplexes, causing them to become sensitive to a minor imperfection such as a single base mispairing. Thus conventional techniques use buffers with low ionic strength, denaturing agents (formamide), or elevated temperatures (usually 50−60 °C) with precise temperature control in order to distinguish the fully matched hybrids from the mismatched ones. However, these conditions do not always lead to the desirable selectivity especially if a mismatch is located at the ultimate or penultimate position of the probe−analyte hybrid.[2] An alternative approach for duplex destabilization has been realized in the case of conformationally constrained probes such as molecular beacons (MBs).[3−5] MBs are oligodeoxyribonucleotide hairpins with a fluorophore and a quencher conjugated to opposite ends of the oligonucleotide. Binding to complementary nucleic acids switches MBs to the elongated conformation and increases their fluorescence. MBs distinguish mismatches over a wider temperature range than unstructured probes do because the stem−loop structure stabilizes the probe−analyte's dissociated state.[4] It was reported that MBs are able to recognize a single substitution in the middle position of 15-mer oligodeoxyribonucleotides.[3,4]

At the same time dividing the probe into two parts leads to an extraordinary selectivity of both protein[6] and nucleic acid recognition.[7] In the case of nucleic acids, each fragment of a binary probe binds to a relatively short (7−10 nucleotide) analyte fragment; this makes each hybrid extremely sensitive to a single base substitution even at mild conditions. Recently one example of such a probe based on malachite green aptamer was reported.[7b] The binary malachite green aptamer probe reliably discriminated against 41 out of 42 possible single nucleotide substitutions in a 14-mer DNA analyte at room temperature in physiological buffer. In the present study, a conformational constraint approach was combined with a "binary approach" in a binary DNA probe (BDP). It was shown that BDP demonstrated an unprecedented high selectivity of DNA recognition: single nucleotide substitutions at any position of a 20 nucleotide DNA analyte could be reliably discriminated at room temperature. The probe was able to detect the true analyte in the presence of a 200-fold excess of a single-base mismatched oligonucleotide.

Materials and Methods

DNAse/RNAse free water was purchased from ICN (Costa Mesa, CA) and used for all buffers and for stock solutions of oligonucleotides. Oligonucleotides were custom-made by Integrated DNA Technologies, Inc. (Coralville, IA) and were used as received. Fluorescent spectra were taken on a Perkin-Elmer (San Jose, CA) LS-55 luminescence spectrometer with a Hamamatsu xenon lamp. Experiments were performed at an excitation wavelength of 485 nm and emission scan of 500−550 nm. The emitting intensities at 517 nM were taken for the calculation of the discrimination factors. The data of four independent experiments were processed using Microsoft Excel.

Discrimination Factors (DFs) for BDPs. The solutions of **MB1** (20 nM), strands **A** and **B** (500 nM for **BDP10/8sl** and **BDP10/8sl F.tul** or 200 nM for **BDP8/8** and **BDP10/8)** in 120 μL of 50 mM

(1) Demidov, V. V.; Frank-Kamenetskii, M. D. *Trends Biochem. Sci.* **2004,** *29*, 62−71.
(2) Urakawa, H.; El Fantroussi, S.; Smidt, H.; Smoot, J. C.; Tribou, E. H.; Kelly, J. J.; Noble, P. A.; Stahl, D. A. *Appl. Environ. Microbiol.* **2003**, *69*, 2848−2856.
(3) Tyagi, S.; Kramer, F. R. *Nat. Biotechnol*. **1996**, *14*, 303−308.
(4) Bonnet, G.; Tyagi, S.; Libchaber, A.; Kramer, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6171−6176.
(5) Marras, S. A.; Tyagi, S.; Kramer, F. R. *Clin. Chim. Acta* **2006**, *363*, 48−60.
(6) (a) Kolpashchikov, D. M. *J. Biomol. Struct. Dyn.* **2003**, *21*, 55−64. (b) Kolpashchikov, D. M.; Rechkunova, N. I.; Dobrikov, M. I.; Khodyreva, S. N.; Lebedeva, N. A.; Lavrik, O. I. *FEBS Lett*. **1999**, *448*, 141−144.
(7) (a) Bichenkova, E. V.; Savage, H. E.; Sardarian, A. R.; Douglas, K. T. *Biochem. Biophys. Res. Commun.* **2005**, *332*, 956−964. (b) Kolpashchikov, D. M. *J. Am. Chem. Soc.* **2005**, *127*, 12442−12443.

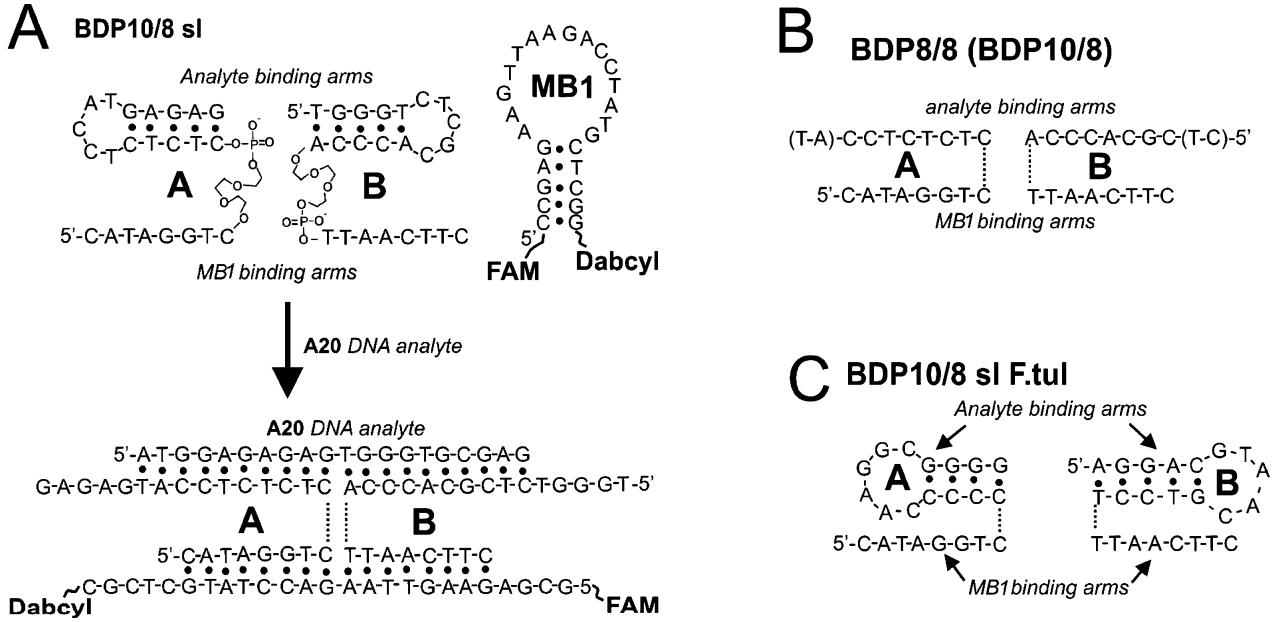

Figure 1. Primary and secondary structure of BDPs used in the present study A: Structure of **BDP10/8sl** in the absence (top) or presence (bottom) of **A20** DNA analyte. B: Structure of strands **A** and **B** of **BDP8/8** and **BDP10/8.** C: Structure of strands **A** and **B** of **BDP10/8sl F.tul.** The triethylene glycole linkers are depicted by the dashed lines on the panels A (bottom), B, and C.

$MgCl_2$, 10 mM Tris HCl, pH 7.4 were incubated in the presence of either 40 nM **A20** or one of the single-base-substituted oligodeoxynucleotide 15 min at room temperature followed by fluorescent emission spectrum measurements. The fluorescence intensities at 517 nM were taken for the calculation of the DFs.

MB2 Assay.[8] Solutions of 20 nM **MB2** (FAM−CTC GCA CCC ACT CTC TCC ATG CGA G−TAMRA) in 120 μL of 100 mM KCl, 1 mM $MgCl_2$, 10 mM Tris HCl were incubated in the presence of either 40 nM **A20** or one of the single base-substituted oligodeoxynucleotides 15 min at room temperature followed by fluorescent emission spectrum measurements. The fluorescence intensities at 517 nM were taken for calculation of the DFs.

A20 Concentration Dependence Measurements. MB1 (20 nM) and strands **A** and **B** (600 nM each) of **BDP10/8sl** were incubated in the absence or presence of 500 nM **A20-4** (ATG *T*AG AGA GTG GGT GCG AG) at variable **A20** concentrations. Fluorescence intensities at 517 nM were taken in four independent experiments.

Results and Discussion

A variant of BDP presented in Figure 1A (**BDP10/8sl)** consists of the two DNA strands **A** and **B** and a molecular beacon (**MB1**). Both strand **A** and **B** comprise the fragments complementary to **MB1** (MB1 binding arms) and the fragments complementary to the nucleic acid analyte (analyte binding arms). Each analyte binding arm contains a structural constraint in the form of a pentanucleotide stem. The analyte binding arm and the MB1 binding arm are connected through triethylene glycole linkers. In the absence of a nucleic acid analyte, the strands are unbound in solution; **MB1** is free in the form of a hairpin (Figure 1A top), and the fluorescence signal is low (Figure 2, curve 1). Addition of **A20** DNA analyte (Figure 1A) triggers the quaternary complex formation (Figure 1A bottom). The fluorophore (FAM) is remote from the quencher (Dabcyl) in this complex, which results in the fluorescence increase (Figure 2, curve 2).

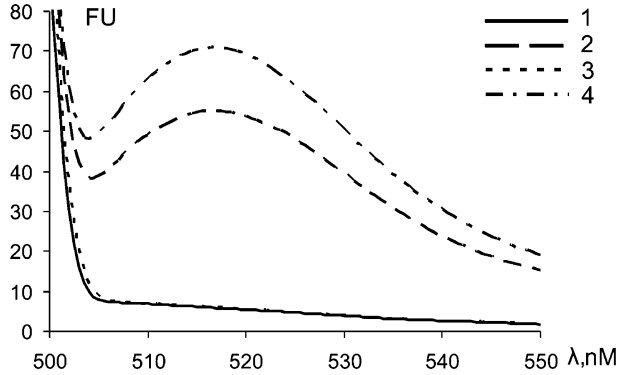

Figure 2. **BDP 10/8sl** increases its fluorescence upon hybridization to **A20** DNA analyte. **MB1** (20 nM) together with strands **A** and **B** (500 nM each) of **BDP 10/8sl** was incubated in the absence (curve 1) or presence (curve 2) of 40 nM **A20**. Curve 3: control assay in the presence of only **MB1** (20 nM). Curve 4: **MB1** in the presence of 40 nM complementary oligodeoxyribonucleotide CAT AGG TCT TAA CTT C.

It was shown that **BDP10/8sl** reliably discriminated **A20** from oligonucleotides containing single-base substitutions at any possible position (Table 1).[9] The discrimination factors (DFs) for **BDP10/8sl** were significantly higher than those of a conventional MB approach. Specifically, **MB2** (FAM−CTC GCA CCC ACT CTC TCC ATG CGA G−Dabcyl), an anti-**A20** molecular beacon, discriminated **A20** from only 15 single-base-substituted oligonucleotides with low DFs (Table 1). Surprisingly, the presence of five substituted oligonucleotides did not trigger **BDP10/8sl** fluorescence significantly above the background (Table 1). Therefore, the probe is insensitive to the presence of **A20-1**, **A20-4**, **A20-8**, **A20-10**, and **A20-17** at the concentration used.

It was interesting, therefore, to test the ability of **BDP10/8sl** to detect the true analyte in solution containing an excess of a

(8) Experiments with **MB2** were carried out in the buffer earlier optimized for selective recognition of DNA analyte by a molecular beacon.[4]

(9) The sequences of the 20 single-base-substituted oligodeoxyribonucleotides were designed in order to introduce one substitution in every position; the type of each substitution was chosen randomly.

Table 1. Discrimination Factors (DFs) for Oligodeoxyribonucleotides Differing from **A20** by a Single Nucleotide[a]

oligodeoxyribonucleotides					
names	sequences	BDP10/8sl	MB2	BDP8/8	BDP10/8
A20	ATGGAGAGAG TGGGTGCGAG	1	1	1	1
A20-1	*T*TGGAGAGAG TGGGTGCGAG	>10	2.1 ± 0.6		>10
A20-2	A*G*GGAGAGAG TGGGTGCGAG	1.8 ± 0.1	0.9 ± 0.2		1.0 ± 0.1
A20-3	ATA*G*AGAGAG TGGGTGCGAG	3.8 ± 1.4	1.1 ± 0.1	1.2 ± 0.1	1.3 ± 0.1
A20-4	ATG*T*AGAGAG TGGGTGCGAG	>10	2.1 ± 0.3	>10	>10
A20-5	ATGG*C*GAGAG TGGGTGCGAG	9.2 ± 1.9	2.3 ± 0.6	>10	5.8 ± 1.2
A20-6	ATGGA*T*AGAG TGGGTGCGAG	9.2 ± 4.1	1.2 ± 0.1	>10	3.5 ± 0.8
A20-7	ATGGAG*G*GAG TGGGTGCGAG	6.9 ± 1.7	1.2 ± 0.2	4.0 ± 0.9	2.1 ± 0.2
A20-8	ATGGAGA*A*AG TGGGTGCGAG	>10	1.6 ± 0.1	>10	>10
A20-9	ATGGAGAG*C*G TGGGTGCGAG	11.5 ± 2.5	1.6 ± 0.1	>10	5.0 ± 0.9
A20-10	ATGGAGAGA*T* TGGGTGCGAG	>10	1.1 ± 0.1	>10	12.6 ± 3.8
A20-11	ATGGAGAGAG *G*GGGTGCGAG	3.0 ± 0.5	1.3 ± 0.2	2.1 ± 0.3	2.2 ± 0.1
A20-12	ATGGAGAGAG T*A*GGTGCGAG	5.5 ± 2.5	1.1 ± 0.1	6.3 ± 1.9	1.9 ± 0.2
A20-13	ATGGAGAGAG TG*T*GTGCGAG	3.5 ± 0.8	1.3 ± 0.1	4.6 ± 1.4	1.6 ± 0.1
A20-14	ATGGAGAGAG TGG*T*TGCGAG	3.5 ± 1.0	1.2 ± 0.1	5.1 ± 1.4	1.8 ± 0.1
A20-15	ATGGAGAGAG TGGG*A*GCGAG	3.4 ± 0.6	1.3 ± 0.2	4.1 ± 1.3	1.6 ± 0.1
A20-16	ATGGAGAGAG TGGGT*T*CGAG	3.2 ± 0.8	1.2 ± 0.1	5.0 ± 1.4	1.8 ± 0.2
A20-17	ATGGAGAGAG TGGGTG*T*GAG	>10	2.5 ± 0.6	10.6 ± 2.2	8.7 ± 1.4
A20-18	ATGGAGAGAG TGGGTGC*C*AG	5.5 ± 1.3	1.3 ± 0.2	2.4 ± 0.4	2.6 ± 0.3
A20-19	ATGGAGAGAG TGGGTGCG*G*G	2.2 ± 0.5	1.2 ± 0.1		2.0 ± 0.3
A20-20	ATGGAGAGAG TGGGTGCGA*A*	1.6 ± 0.3	1.1 ± 0.1		1.7 ± 0.2

[a] Discrimination factors were calculated as the ratio of the BDP or MB2 fluorescence intensity at 517 nM in the presence of **A20** (true target) to the fluorescence intensities in the presence of each mismatched oligonucleotide after subtraction of the background fluorescence. The DFs were estimated as being higher than 10 (the signal-to-background ratio) for those oligonucleotides that triggered no fluorescence significantly above the background. The mismatched positions are underlined. The data are the average of four independent measurements.

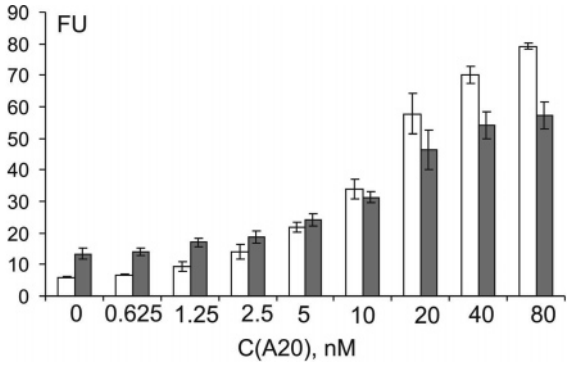

Figure 3. The dependence of **BDP10/8 sl** fluorescence intensity on **A20** concentration in the absence (white bars) or presence (gray bars) of 500 nM **A20-4** (ATG *T*AG AGA GTG GGT GCG AG).

single-base-substituted oligonucleotide. The presence of 500 nM **A20-4** was found to increase the background fluorescence by only about 2 times (Figure 3). At the same time, the background value was significantly exceeded when a 2.5 nM true target was added (Figure 3). In this case, the concentration of the analyte was only 0.5% of the concentration of the single-base-substituted oligonucleotide. This extremely high specificity of the probe can be used in minority point mutation analysis directly after conventional PCR, thus avoiding additional work-intensive enzyme-mediated steps used by more conventional approaches.[10]

It should be noted that a BDP variant containing no structural constraints (Figure 1B) demonstrated excellent selectivity in recognition of a 16 nucleotide DNA fragment (Table 1). However, when the analyte binding arms were elongated to 10 nucleotides, the selectivity was notably decreased (compare discrimination factors in Table 1). These data demonstrate that the combination of conformational constraint with a "binary approach" (splitting probe into two halves) results in a synergistic effect and leads to a significantly improved selectivity of nucleic acid recognition. This effect is stipulated by the cooperative nature of the quaternary complex: it dissociates into four rather than two fragments, leading to a higher entropy gain in comparison to the conventional monolith probe. Moreover, the three DNA stem−loop structures formed upon dissociation provide the compensation of enthalpy lost. Thus both the enthalpy and the entropy contributions in the free energy of the probe/analyte dissociated state enhance the dissociation process, especially in the presence of a mismatch base-pairing.

To confirm the universal character of the suggested approach, a BDP that recognizes an alternative 20 nucleotide DNA analyte was designed. Recently an application of the MB approach was reported for discrimination of a unique region of the 16S rRNA of the bacterium *Francisella tularensis* from its mutants.[11] *F. tularensis* is a potential biological weapon agent and a highly infectious bacterium. Poor discrimination was found for single-base-substituted **CS111** and double-base-substituted **CS120** and **CS122** (for structures see Table 2) even in the optimized buffers.[11]

The structure of strands **A** and **B** of the BDP were designed for the recognition of the *F. tularensis* sRNA fragment represented by oligodeoxyribonucleotide **CS11**. **BDP10/8sl F.tul** (Figure 1C) contained MB1 binding arms of the same structure as that of previously described BDPs. On the other hand the analyte binding arms were complementary to **CS11** and contained four-nucleotide (strand **A**) or five-nucleotide (strand **B**) stems. It was found that **BDP10/8sl F.tul** reliably discriminated all oligonucleotides tested from the true analyte (Table 2). It should be emphasized that 4 out of 10 single-base-substituted oligonucleotides including **CS120** and **CS122** triggered no fluorescence substantially higher than the background;

(10) Makrigiorgos, G. M. *Human Mut.* **2004**, *23*, 406−412.

(11) Ramachandran, A.; Flinchbaugh, J.; Ayoubi, P.; Olah, G. A.; Malayer, J. R. *Biosens. Bioelectron.* **2004**, *19*, 727−736.

Table 2. Discrimination Factors for **BDP10/8 F.tul**[a]

oligonucleotide		
name	sequence	DF
CS11	GCCTTGGGGG AGGACGTTAC (true target)	1
CS111	GC*T*TTGGGGG AGGACGTTAC	1.6 ± 0.2
CS112	*A*CCTTGGGGG AGGACGTTAC	1.6 ± 0.1
CS113	G*T*CTTGGGGG AGGACGTTAC	>5.0
CS114	GCCTTGGGG*A* AGGACGTTAC	10.1 ± 1.8
CS115	GCCTTGGGGG AGGACGTTA*T*	7.1 ± 0.7
CS116	GCCTTGGGGG AG GA*T* GTTAC	2.2 ± 0.3
CS117	GCCTTGGGGG AGGACGT*C*AC	>5.0
CS120	GCCTTGGGG*A G*GGACGTTAC	>5.0
CS122	G*T*CTTGGGG*A* AGGACGTTAC	>5.0

[a] DFs were calculated as described in the footnote to Table 1. The data are the average of four independent experiments. The DFs were estimated as being higher than 5 (signal-to-background ratio) for those oligonucleotides that triggered no fluorescence significantly above the background.

this result demonstrates a remarkably improved selectivity of target recognition in comparison to the MB approach.[1]

The design of BDPs was straightforward and easy: the sequence of the analyte binding arms was the only change made for BDP adjustment to the new analyte, while the same MB was used as a fluorescent reporter. Taking into account the relative expense of MB probes, the BDP may substantially reduce the cost of multitarget assays in comparison to the conventional MB approach.

Possibly, the simplest construction of a BDP would be the one that forms the classical DNA four-way junction (Holiday junction)[12] in the presence of the analyte. In this case, out of four strands involved in the complex formation, one is an analyte, the second is an MB, and the two others are unmodified DNA strands. However, it was shown earlier that Holiday junctions exist as a mixture of two right-handed antiparallel crosses.[12] Only one of these conformers contains an MB in an elongated form allowing a substantial fluorescence increase. In general, the ratio of the two conformers depends on the nucleotide composition at the point of strand exchange.[12] In the case of BDP, this is predetermined by the analyte sequence. In practice it is convenient to make MB in an elongated conformation with analytes of any sequences. This is realized in the BDPs described in this study. In the quaternary complex (Figure 1A bottom), MB and the analyte preferably acquire elongated conformations, while strands **A** and **B** are bent at the point of the oligoethylene glycol inserts.

The Holiday junction-like structure formed in the presence of DNA analyte contains a branch point and, therefore, is a motif useful for bottom-up building of DNA nanostructures.[13] In this respect, BDP can be considered as an initial application of structural DNA nanotechnology in nucleic acid analysis.

Acknowledgment. The author is grateful to Milan N. Stojanovic for support of this work. I also thank Olga N. Ozoline and Yuliya V. Gerasimova for helpful discussions. This study was funded by the NIH, NIBIB, RO1 EB000675 and NSF, BES 0321972 (PI M. N. S.)

JA0628093

(12) Lilley, D. M. J. *Q. Rev. Biophys*. **2000**, *33*, 109−159.
(13) Seeman, N. C. *Q. Rev. Biophys.* **2006**, 1−9.