A Differential Fluorescent Receptor for Nucleic Acid Analysis

Hillary N. Bengtson and **Dmitry M. Kolpashchikov**[*]
Chemistry Department and Burnett School of Biomedical Sciences University of Central Florida
4000 Central Florida Blvd, Orlando, FL 32816 (USA)

Keywords

differential receptor; SNP analysis; dumbbell molecular beacon; principal component analysis

Human senses of taste and smell inspired the development of artificial differential receptors for the recognition and identification of a great variety of chemical species.[1] A differential receptor consists of an array of sensors. Each sensor in the array recognizes a series of analytes with different recognition rates. The receptor produces a pattern of the responses. Analysis of the pattern (the fingerprint of the analyte) reveals the presence of a particular analyte in the sample. For example, Anslyn and colleagues designed a differential receptor for the recognition of fatty acids based on serum albumins.[2] Stojanovich et al. used DNA three-way junction sensors to design a differential receptor for steroids.[3] Recently, Chou et al. used nanoscale graphene oxide to design a receptor that differentiated proteins.[4] Here, we adopt this concept for differential analysis of nucleic acid sequences.

One of the most challenging tasks in nucleic acid analysis is to distinguish two related sequences that differ by a single nucleotide; such as, single nucleotide polymorphisms (SNPs) or point mutations.[5] A conventional approach for SNP analysis takes advantage of a fluorescent sensor called the molecular beacon (MB) probe (Figure 1A).[6] It is a stem-loop folded oligonucleotide conjugated with a fluorophore dye at the 5' end and with a quencher dye at the 3' end. The stem brings the fluorophore close to the quencher, while the loop fragment is complementary to an analyzed SNP site. When the MB probe hybridizes with a fully complementary (matched) analyte it undergoes a conformational change to the elongated form, which results in brighter fluorescence (Figure 1B). According to the conventional paradigm, the length of the loop portion and the hybridization conditions are selected to minimize binding of the MB probe to a mismatched analyte, so that fluorescent signal remains at the background level. This approach corresponds to YES/NO digital response. Therefore, analysis of each SNP site requires the synthesis and optimization of two MB probes to report each of two alleles. For genotyping hundreds and thousands of SNP sites double the amount of expensive MB probes are required.

Recently, El-Hajj et al. introduced sloppy MB probes[7a] that are SNP tolerant and can be used for genotyping of SNPs in a differential receptor format.[7] However, this approach

---

[*]dmk2111@gmail.com.

requires a number of expensive MB probes, each of which should undergo costly optimization for accurate SNP differentiation. In this study we introduce a differential fluorescent receptor (DFR) that requires only one fluorescent reporter to differentiate at least eight and potentially a greater number of analytes. The overall idea is to create an array of sensors that produce a characteristic SNP-specific signature in the presence of each individual nucleic acid sequence. Each DFR is based on triple-crossover (TX) sensors that use a dumbbell molecular beacon (DMB) probe reported by Lv et al.[8] (Figure 1C, top left). We turned our attention to this MB probe variation, since the DFR concept requires a mutation-tolerant reporter that not only fluoresces in complex with fully matched analyte, but also produces signals of different (presumably lower) intensities in the presence of mismatched sequences. A DMB contains two stem-loops that stabilize the dumbbell shape. Overall this design allows the combination of a relatively long probe sequence with stable secondary structure. In addition, it avoids long stems, which reduce the hybridization rates.[9] Therefore, DMB represents an attractive alternative to the conventional MB probe if a SNP-tolerant nucleic acid recognition is required. Since double-labeled fluorescent probes are relatively expensive, we designed an array of fluorescent sensors that utilizes a single DMB probe as a universal reporter.

Each TX sensor utilizes a universal DMB probe and four target-specific adaptor strands (A, B, C and D), as illustrated in Figure 1C. Each adaptor strand contains a fragment complementary to the analyte and one or two fragments complementary to the DMB probe. The DMB probe and the adaptor strands can form a hexapartite TX complex with the fully complementary analyte (Figure 1C, bottom). In this complex, the fluorophore is separated from the quencher, which enables bright fluorescence. The complex contains three DNA four-way junctions, which resemble the crossover tiles characterized earlier by structural DNA nanotechnology.[10] Presence of mismatches between one of the adaptor strands and the analyte may result in a less stable complex, thus shifting hybridization equilibrium towards free DMB probe. This creates a basis for the differential response to the presence of nucleic acid sequences containing a variety of SNPs. In this proof-of-concept study we demonstrated the feasibility of the aforementioned design.

We chose the sequences of the *rpoB* gene of Mycobacterium tuberculosis (*M.tb*) as a clinically relevant analyte. *M.tb* infects approximately 2 billion people worldwide and is responsible for about 2 million deaths each year.[11] Approximately 10% of all patients are infected with drug-resistant *M.tb* strains, mainly to the antibiotics rifampin and isoniazid.[12] Rapid and accurate drug susceptibility testing is crucial for the control of active TB with effective drug cocktails. We chose a wild type (WT) and seven sequences of rifampin-resistant strains that have been studied and characterized previously.[13] All the mutations were located in a characteristic 81-bp fragment of the *rpoB* gene. This fragment is known to contain 95 to 98% of all mutations that impart rifampin resistance to *M.tb*.[13] We designed three fluorescent sensors, with a single universal MB probe reporter, targeting different segments of the 81-bp fragment of the *rpoB* gene. Each TX sensor hybridizes to a different region of the analyte due to the unique set of the adaptor stands.

Sensor 1 used adaptor strands A1, B1, C1, and D1 (Figure 2A and S1). The DMB-binding arms of the strands were designed short to minimize interaction with DMB1 probe in the

absence of analytes. In particular, strands B1 and C1 contained two pentaor hexa-nucleotide DMB binding arms. In the presence of a cognate DNA analyte, the four adaptor strands and DMB1 probe formed a hexapartite complex with the analyte (Figure 2A, bottom). The fluorophore group is remote from the quencher in this complex, which instigates high fluorescence. Indeed, the complementary DNA analyte, MT6, triggered ~ 7-fold fluorescence increase (Figures 2B and S2). However, when mismatched nucleotides were present, the stability of the complex was decreased for most of the analytes (except MT1) resulting in the reduced fluorescent signal. The higher signal for MT1 analyte can be explained after analysis of secondary structures of the analytes (see SI, Figure S1 and accompanying discussion). Overall, Sensor 1 differentiated WT from MT1, MT5 and MT6. However, in the presence of WT, the sensor signal was statistically indistinguishable from that in the presence of MT2, MT3, MT4 and MT7. To enable differentiation of all analytes, we introduced additional sensors according to the paradigm of differential receptors.[1] The structures and fluorescent responses of Sensors 2 and 3 are presented in Table S1, Figures S3 and S4. The sensors were designed to recognize different fragments of the analytes. Each sensor produced individual pattern of fluorescent responses (Figures S2 and S3). For example, Sensor 2 differentiated WT from MT1, MT2, MT3, and MT4 (Figure S2), while Sensor 3 differentiated WT from MT4, MT5, MT6 and MT7 (Figure S4). Time required for fluorescent responce of the TX sensors was similar to that for a conventional MB probe[6] and other MB probe-based sensors[14] (Figure S5).

The fluorescent data obtained from the three DFR sensors was analyzed using principal component analysis (PCA),[15] a chemometric technique, that is conventionally used to interpret multidimensional sets of data including those generated by differential receptors.[1] The PCA score plot shown in Figure 3 demonstrates that all eight analytes are distinguished. MT3 and MT4 contained mutations at the same position; therefore, clustering of responses in the close areas was expected for these two analytes. The ability of the sensor to differentiate the analytes at different concentrations was investigated (Figure S6). It was found that DFR was able to distinguish most analytes. The differentiation power of the receptor can be further increased by optimization of individual TX sensors or by introducing additional sensors to the array. However, even with the current differentiaon rates the DFR might be useful in practice.

For example, Figure 4 represents an alternative analysis of the data set, in which fluorescent intensities of each series were divided by those of WT (F/F_{WT}). Although each sensor failed to differentiate all eight analytes from WT, an array of only three sensors was sufficient for differentiation of all 8 altered DNA sequences from the WT. In this representation, the signal pattern obtained for WT significantly differed from that of the mutants MT1-7. Therefore, the DMB-based differential receptor is promising for the distinction of drug-susceptible *M.tb* species from the drug-resistant ones, which is of great practical significance.[12,13]

The differential receptor concept was inspired by the natural senses of smell and taste and has been adopted for sensing of a broad spectrum of biological molecules.[1-4] Here, we propose to use a differential receptor to identify a series of clinically important DNA sequences that differ by a single nucleotide. We demonstrated that an array of three TX

sensors that utilize the same DMB fluorescent reporter is sufficient to differentiate eight analytes. The design of the sensors was straightforward; the analyte-binding arms were the only change in the design of each new sensor. The use of non-traditional DMB probe as a fluorescent signal reporter is justified by the need for the increased stability of the folded MB state in the presence of four adaptor strands. The sensor is potentially compatible with such practically significant formats as quantitative real-time PCR (qPCR). Indeed, same as in MB probe-based qPCR, each TX sensor can potentially serve as a real-time hybridization probe. Importantly, the sensor can be adjusted to operate at elevated temperatures (45-55°C) used in qPCR by changing the length of the adaptor strands. We plan to explore this possibility in our future work. An array of several (three) PCR assays utilizing different TX sensor would produce a pattern of fluorescent signals which serve as a unique signature of each analyte. Differential receptors designed against small molecules are typically challenged by the problem of differentiating analytes in complex mixtures of related compounds. This problem is less relevant to qPCR-based analysis of bacterial species in clinical samples, as typically only one type of pathogen is assayed with a given set of PCR primers and the co-infection with two species of bacteria is rear.

In conclusion, we have introduced a fluorescent sensor for the analysis of single nucleotide substitutions in DNA. With this report we hope that the differential receptor concept will find applications in diagnosis of infectious diseases, genetic disorders and in genetic linkage analysis.

Experimental Section

All oligonucleotides were custom-made by Integrated DNA Technologies, Inc (Coralville, IA). The fluorescent assay contained the DMB1 probe (40 nM), each of the four adaptor strands A-D at concentrations 0.4-1.2 μM (see Table S1), and one of the analytes WT, or MT1-MT7 (100 nM). Final sample volumes were 120 μL. All the components of the assay were mixed in the buffer containing 50 mM Tris-HCl, pH 7.4, 50 mM $MgCl_2$ and incubated at 22°C for 25 min. Then, fluorescence of the samples was measured on a Perkin-Elmer (San Jose, CA) LS-55 Luminescence Spectrometer with a Hamamatsu xenon lamp (excitation at 485 nm; emission at 517 nm). The data was analyzed using Microsoft Excel. Principal component analysis was done using MeV v4.8.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. a Liu Y, Ding LP, Cao Y, Fang Y. Prog. Chem. 2012; 24:1915–1927.b Wright AT, Anslyn EV. Chem. Soc. Rev. 2006; 35:14–28. [PubMed: 16365639] b Albert KJ, Lewis NS, Schauer CL,

Sotzing GA, Stitzel SE, Vaid TP, Walt DR. Chem. Rev. 2000; 100:2595–2626. [PubMed: 11749297] c Lavigne JJ, Anslyn EV. Angew. Chem. Int. Ed. 2001; 40:3119–3130.j Folmer-Andersen JF, Kitamura M, Anslyn EV. J. Am. Chem. Soc. 2006; 128:5652–5653. [PubMed: 16637629] k Scott SM, James D, Ali Z. Microchim. Acta. 2006; 156:183–207.m Zhou H, Baldini L, Hong J, Wilson AJ, Hamilton AD. J. Am. Chem. Soc. 2006; 128:2421–2425. [PubMed: 16478197] m Collins BE, Wright AT, Anslyn EV. Top. Curr. Chem. 2007; 277:181–218.n Palacios MA, Wang Z, Montes VA, Zyryanov GV, Anzenbacher P Jr. J. Am. Chem. Soc. 2008; 130:10307–10314. [PubMed: 18616249] o Röck F, Barsan N, Weimar U. Chem. Rev. 2008; 108:705–725. [PubMed: 18205411] p Phaisangittisagula E, Nagle HT, Areekul V. Sens. Actuators B. 2010; 145:507–515.q Rochat S, Gao J, Qian X, Zaubitzer F, Severin K. Chem. Eur. J. 2010; 16:104–113. [PubMed: 19938007] r Umali AP, Anslyn EV. Curr. Opin. Chem. Biol. 2010; 14:685–692. [PubMed: 20801075] s Carey JR, Suslick KS, Hulkower KI, Imlay JA, Imlay KRC, Ingison CK, Ponder JB, Sen A, Wittrig AE. J. Am. Chem. Soc. 2011; 133:7571–7576. [PubMed: 21524080] t Stewart S, Syrett A, Pothukuchy A, Bhadra S, Ellington A, Anslyn E. ChemBioChem. 2011; 12:2021–2024. [PubMed: 21796750]

2. Kubarych CJ, Adams MM, Anslyn EV. Org Lett. 2010; 12:4780–4783. [PubMed: 20942411]

3. a Stojanovi MN, Green EG, Semova S, Niki DB, Landry DW. J Am. Chem. Soc. 2003; 125:6085–6089. [PubMed: 12785838] b Yang KA, Pei R, Stefanovic D, Stojanovic MN. J. Am. Chem. Soc. 2012; 134:1642–1647. [PubMed: 22142383] c Green E, Olah MJ, Abramova T, Williams LR, Stefanovic D, Worgall T, Stojanovic MN. J. Am. Chem. Soc. 2006; 128:15278–15282. [PubMed: 17117880] d Pei R, Shen A, Olah MJ, Stefanovic D, Worgall T, Stojanovic MN. Chem Commun (Camb). 2009; 22:3193–3195. [PubMed: 19587910]

4. Chou SS, De M, Luo J, Rotello VM, Huang J, Dravid VP. J. Am. Chem. Soc. 2012; 134:16725–16733. [PubMed: 22962967]

5. a Bedeir A, Krasinskas AM. Arch. Pathol. Lab. Med. 2011; 135:578–587. [PubMed: 21526956] b Schumacher JA, Szankasi P, Bahler DW, Ho AK, Kelley TW. J. Clin. Pathol. 2011; 64:618–625. [PubMed: 21486895] c Chiou KR, Charng MJ, Chang HM. Atherosclerosis. 2011; 216:383–389. [PubMed: 21376320] d Wang Y, Cottman M, Schiffman JD. Cancer Genet. 2012; 205:341–355. [PubMed: 22867995] e Nepomnyashchaya YN, Artemov AV, Roumiantsev SA, Roumyantsev AG, Zhavoronkov A. Clin. Chem. Lab Med. 2012; 29:1–14.f Ahmad A, Iqbal MA. Curr. Med. Chem. 2012; 19:3739–3747. [PubMed: 22680919]

6. a Tyagi S, Kramer FR. Nature Biotech. 1996; 14:303–308.b Huang K, Martí AA. Anal. Bioanal. Chem. 2012; 402:3091–3102. [PubMed: 22159461] c Li J, Cao ZC, Tang Z, Wang K, Tan W. Methods Mol Biol. 2008; 429:209–224. [PubMed: 18695969] d Kolpashchikov DM. Scientifica. 2012:ID 928783. 2012.

7. a El-Hajj HH, Marras SA, Tyagi S, Shashkina E, Kamboj M, Kiehn TE, Glickman MS, Kramer FR, Alland D. J Clin Microbiol. 2009; 47:1190–1198. [PubMed: 19171684] b Chakravorty S, et al. J Clin Microbiol. 2012; 50:2194–2202. [PubMed: 22535987]

8. Lv C, Yu L, Wang J, Tang X. Bioorg. Med. Chem. Lett. 2010; 20:6547–6550. [PubMed: 20933418]

9. a Dirks RM, Pierce NA. Proc. Natl. Acad. Sci. U S A. 2004; 101:15275–15278. [PubMed: 15492210] b Tsourkas A, Behlke MA, Rose SD, Bao G. Nucleic Acids Res. 2003; 31:1319–1330. [PubMed: 12582252] d Kolpashchikov DM. Chembiochem. 2009; 10:1443–1445. [PubMed: 19444831]

10. a Fu TJ, Seeman NC. Biochemistry. 1993; 32:3211–3220. [PubMed: 8461289] b Kim S, Kim J, Qian P, Shin J, Amin R, Ahn SJ, Labean TH, Kim MK, Park SH. Nanotechnology. 2011; 22:245706. [PubMed: 21543827] c Sa-Ardyen P, Vologodskii AV, Seeman NC. Biophys J. 2003; 84:3829–3837. [PubMed: 12770888] d Kolpashchikov DM, Gerasimova YV, Khan MS. Chembiochem. 2011; 12:2564–2567. [PubMed: 22006680]

11. a Pai M, Ramsay A, O'Brien R. PLoS Med. 2008; 5:e156. [PubMed: 18651788] b bu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM Jr. et al. Proc Natl Acad Sci U S A. 2009; 106:13980–13985. [PubMed: 19666590] c Paramasivan CN, Lee E, Kao K, et al. Int. J. Tuberc. Lung Dis. 2010; 14:59–64. [PubMed: 20003696]

12. a WHO Anti-Tuberculosis Drug Resistance in the World. 2008b WHO Multidrug and Extensively Drug-Resistant TB (M/XDR-TB). Global Report on Surveillance and Response. 2010

13. a Telenti A, Imboden P, Marchesi F, Lowrie D, Cole S, Colston MJ, Matter L, Schopfer K, Bodmer T. Lancet. 1993; 341:647–650. [PubMed: 8095569] b Donnabella V, Martiniuk F, Kinney D, Bacerdo M, Bonk S, Hanna B, Rom WN. Am. J. Respir. Cell Mol. Biol. 1994; 11:639–643. [PubMed: 7946393] c Morris S, Bai GH, Suffys P, Portillo-Gomez L, Fairchok M, Rouse D. J. Infect. Dis. 1995; 171:954–960. [PubMed: 7706824] d Helb D, et al. J. Clin. Microbiol. 2010; 48:229–237. [PubMed: 19864480] e El-Hajj HH, Marras SA, Tyagi S, Kramer FR, and D. All. J. Clin. Microbiol. 2001; 39:4131–4137. [PubMed: 11682541] f Cooksey RC, Morlock GP, Glickman S, Crawford JT. J. Clin. Microbiol. 1997; 35:1281–1283. [PubMed: 9114427] g Troesch A, Nguyen H, Miyada CG. J. Clin. Microbiol. 1999; 37:49–55. [PubMed: 9854063] h Gingeras TR, et al. Genome Res. 1998; 8:435–448. [PubMed: 9582189] i Blakemore R, Story E, Helb D, Kop J, Banada P, Owens MR, Chakravorty S, Jones M, Alland D. Clin Microbiol. 2010; 48:2495–2501.j Miotto P, Bigoni S, Migliori GB, Matteelli A, Cirillo DM. Eur. Respir. J. 2012; 39:1269–1271. [PubMed: 22547737] k Armand S, Vanhuls P, Delcroix G, Courcol R, Lemaître N. J. Clin. Microbiol. 2011; 49:1772–1776. [PubMed: 21411592] l Malbruny B, Le Marrec G, Courageux K, Leclercq R, Cattoir V. Int. J. Tuberc. Lung Dis. 2011; 15:553–555. [PubMed: 21396219] m Vadwai V, Boehme C, Nabeta P, Shetty A, Alland D, Rodrigues C. J. Clin. Microbiol. 2011; 49:2540–2545. [PubMed: 21593262] n Gous N, Scott LE, Wong E, Omar T, Venter WD, Stevens W. Clin. Microbiol. 2012; 50:2100–2103.

14. a Cornett EM, O'steen MR, Kolpashchikov DM. PLOS one. 2013; 8:e55919. [PubMed: 23441157] b Gerasimova YV, Kolpashchikov DM. Biosensors & Bioelectronics. 2013; 41:386–390. [PubMed: 23021850]

15. a Abdi H, Williams LJ. Wiley Interdisciplinary Reviews: Computational Statistics. 210:2, 433–459.b Shaw PJA. Multivariate statistics for the Environmental Sciences, Hodder-Arnold. 2003
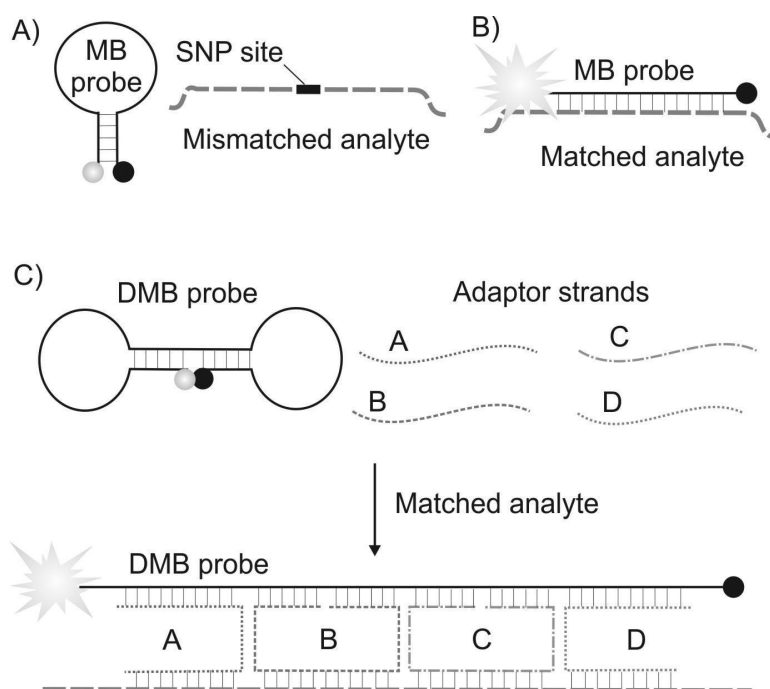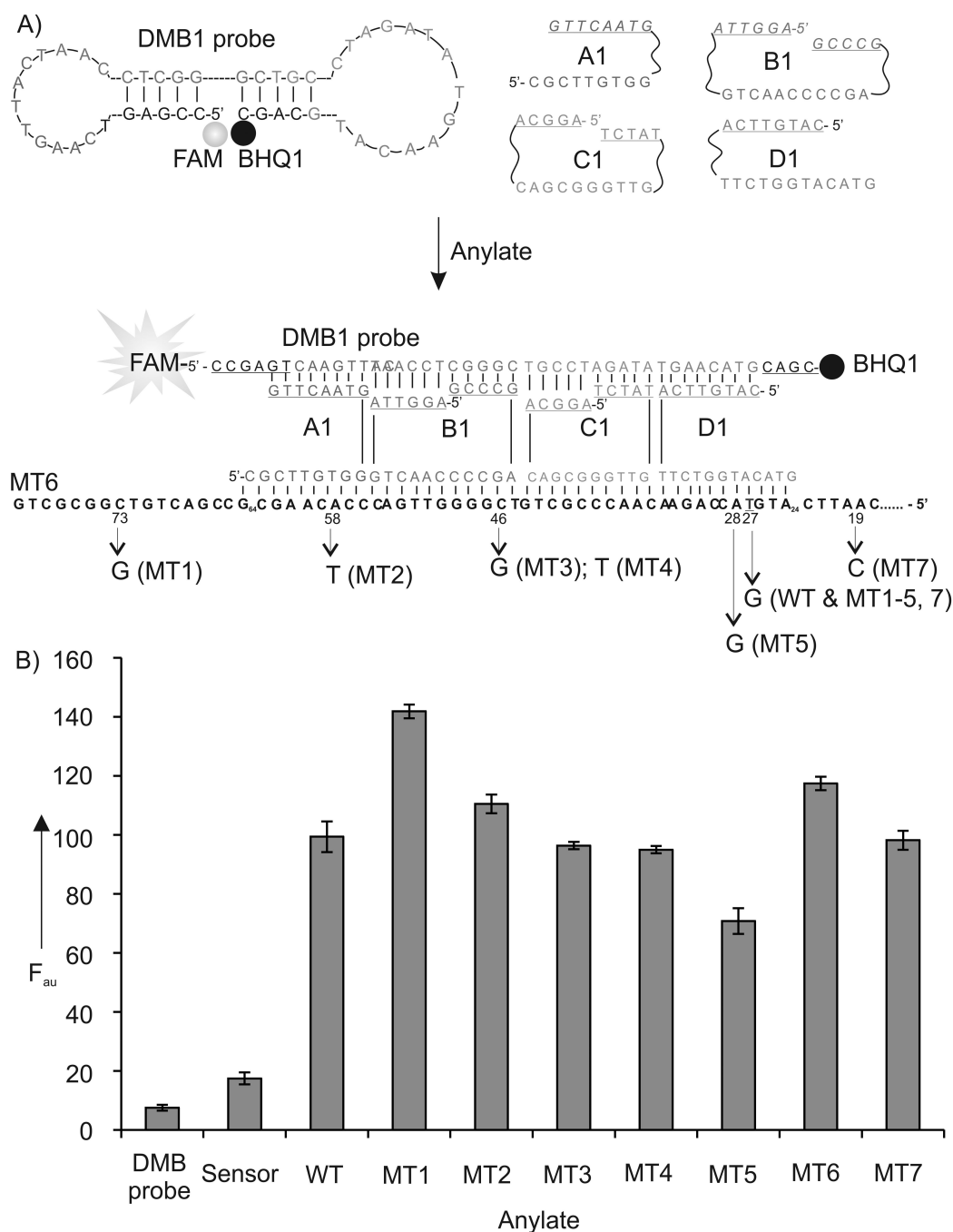
Figure 1.
Comparison of the molecular beacon (MB) probe and differential fluorescent receptor (DFR) approaches for the analysis of SNP containing nucleic acids. A) MB probe does not hybridize to mismatched DNA sequence (the analytes are shown as dashed lines). B) MB probe hybridizes to matched DNA analyte. C) DFR consists of a dumbbell molecular beacon (DMB) probe and four adaptor strands. DFR forms a highly fluorescent complex with a fully matched analyte.

Figure 2.
The structure and performance of Sensor 1. A) Schematic diagram of Sensor 1 in the absence and in the presence of MT6 analyte. DMB-binding arms of adaptor strands A1, B1, C1 and D1 are underlined. Single nucleotide differences in DNA analytes are indicated on the bottom. FAM is fluorescein; BHQ1 is Black Hole Quencher-1. B) Fluorescence of Sensor 1 in the presence of different analytes. Samples containing DMB probe (40 nM), A1 (1200 nM), B1, C1, D1 (800 nM each), and analytes (100 nM each) were incubated in the buffer containing 50 mM Tris-HCl, pH 7.4, 50 mM MgCl$_2$ for 25 minutes at 22°C followed

by the measurement of FAM fluorescence at 517 nm upon excitation at 485 nm.The data are average values of three independent trials with standard deviations.
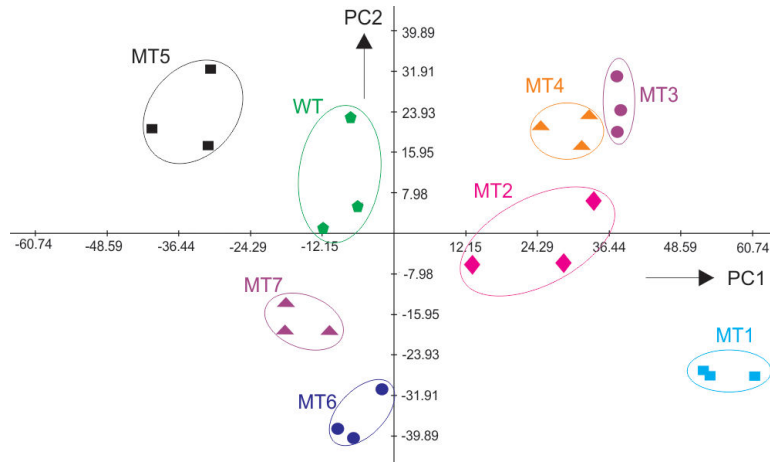
Figure 3.
Principal component analysis (PCA) score plot of the responses from different *M.tb* analytes in the differential fluorescent receptor (DFR) assay. Clustering of the signals for each analyte in non-overlapping areas indicates the ability of the DFR to differentiate all eight DNA analytes.
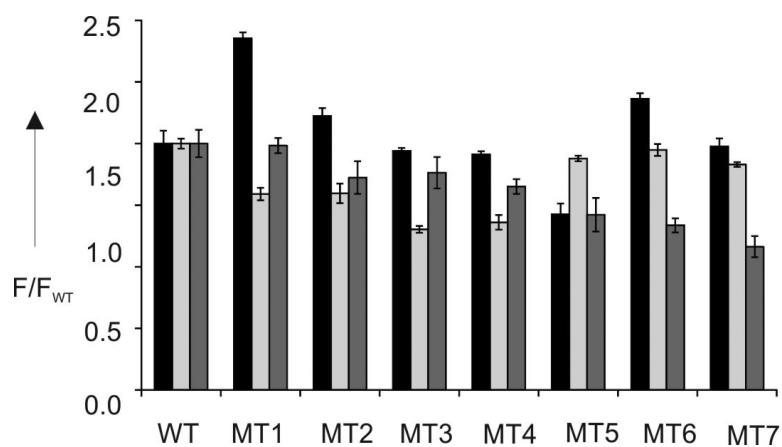
Figure 4.
Fluorescent signatures of eight *M.tb* DNA analytes produced by the differential fluorescent receptor (DFR) assay. The black, light grey and dark grey bars represent fluorescent signals of Sensors 1, 2, and 3, respectively, divided by the average florescence produced by WT analyte (F/F_{WT}). The data of three independent measurements are shown with standard deviations.