Design of a novel molecular beacon: modification of the stem with artificially genetic alphabet†

Pinpin Sheng,[a] Zunyi Yang,*[b] Youngmi Kim,[a] Yanrong Wu,[a] Weihong Tan*[a] and Steven A. Benner*[b]

A molecular beacon that incorporates components of an artificially expanded genetic information system (AEGIS) in its stem is shown not to be opened by unwanted stem invasion by adventitious standard DNA; this should improve the "darkness" of the beacon in real-world applications.

The desire to detect specific DNA and RNA sequences quantitatively and with high throughput has created a demand for biomolecular recognition probes that respond with high signal-to-noise ratios, low background, and few false positives in complex biological environments. These specifications are needed especially for molecular beacons (MBs), which are short oligonucleotide probes that carry a fluorescent molecule on one end and a quencher on the other. These come in close contact in a hairpin structure (Fig. 1(a)), but become separated when a target oligonucleotide analyte binds to their loop regions.[1] Once the beacon is opened, its fluor can generate a fluorescence signal essentially free of quenching.

Background fluorescence arises in MBs from various sources. First, quenching is never perfect, even when the fluor and quencher are in close proximity. Further, a class of false positives may arise when the beacon is digested with nucleases,[2a] or opened by binding to a protein.[2b] These issues may be managed using nuclease-resistant backbones (e.g. phosphorothioates or 2′-O-methyl RNA sugars),[3a,b] or by using sugars not recognized by DNA- and RNA-binding proteins (e.g. PNA,[4a,b] LNA,[4c] or HOMO DNA[4d]).

Another source of false positives arises when the MB has access in thermodynamic equilibrium to a different, non-hairpin intramolecular structure that pairs one of the stem segments with a segment in the loop (a conformational switch) (Fig. 1(b)); this generates background to the extent that the alternative conformation is present at equilibrium, and the distance is long between the fluor and the quencher in that conformation.[5] While this might be managed by careful design of the beacon structure, rules that guide this design are imperfect, nearly always requiring a "winnowing" of alternative designs. False positives also arise if the MB is opened by binding to a standard oligonucleotide not in its loop region, but rather in its stem region (with perhaps some binding in the loop as well) (Fig. 1(c)). While this problem may be overcome using L-DNA in the stem,[6] L-DNA phosphoramidites (Chem-Genes Corporation, Wilmington, MA) are quite expensive.

Accordingly, we examined an alternative approach to design MBs that relies less on design. This approach relies on an artificially expanded genetic information system (AEGIS) (Fig. 2), which was developed some time ago in these laboratories.[7] Pairing between AEGIS components in a duplex retains an overall Watson–Crick geometry, and involves three hydrogen bonds. AEGIS pairs therefore contribute to duplex stability to the same (or greater) extent as standard pairs. Because of non-standard arrangement of their hydrogen bonding functionality, however, AEGIS components should not bind to standard nucleotides. Therefore, placing AEGIS components into the stem of a MB should eliminate any opening of a beacon by adventitious natural DNA.

To test this concept, we exploited 6-amino-5-nitro-3-(1′-beta-D-2′-deoxyribofuranosyl)-2(1H)-pyridone (d**Z**), and its complement 2-amino-8-(1′-beta-D-2′-deoxyribofuranosyl)-imidazo[1,2-a]-1,3,5-triazin-4(8H)-one (d**P**) as the AEGIS pair.[8] This has advantages with respect to chemical stability, enzymatic compatibility, and cost over another AEGIS pair,
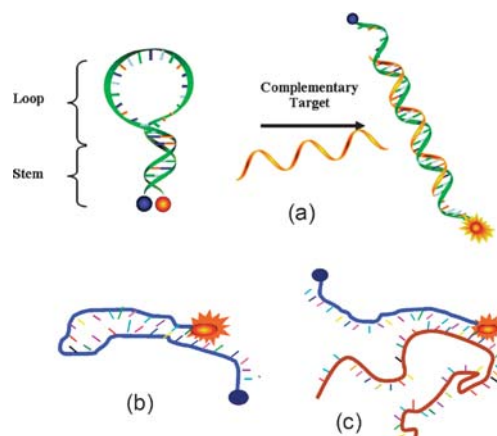

Fig. 1 Molecular beacon design and interactions. (a) MB binding to complementary target; (b) MB mis-folds into non-hairpin structure through the hybridization of loop sequence with stem sequence; (c) MB opened by adventitious DNA that binds to MB stem.

[a] Department of Chemistry and UF Genetics Institute, Shands Cancer Center, Center for Research at Bio/nano Interface and McKnight Brain Institute, University of Florida, Gainesville, FL 32611-7200, USA. E-mail: tan@chem.ufl.edu; Fax: 1 352 846 2410; Tel: 1 352 846 2410
[b] Foundation for Applied Molecular Evolution, 1115 NW 4th Street, Gainesville, FL 32601, USA. E-mail: sbenner@ffame.org; zyang@ffame.org; Fax: 1 352 271 7076; Tel: 1 352 271 7005
† The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
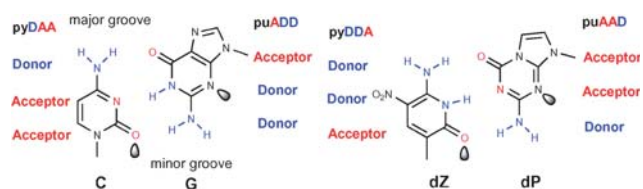
Fig. 2 One example of an "artificially expanded genetic information system" (AEGIS). Nucleobase pairs in this system have a Watson–Crick geometry, with large purines or purine analogs (indicated by "pu") pairing with small pyrimidines or pyrimidine analogs (indicated by "py") joined by hydrogen bonds. The hydrogen-bonding acceptor (A) and donor (D) groups are listed from the major to the minor groove as indicated. Lobes represent electron density in the minor groove. The nucleotides implementing the pyDDA:puAAD hydrogen bonding pattern (dZ:dP), the topic of this paper, are at the right.

5-methylisocytidine (diso^{Me}C) and isoguanosine (disoG), widely used today in FDA approved diagnostic assays that help manage the care of some 400 000 patients infected with HIV, hepatitis B and hepatitis C viruses.[9] Hybridization of the dZ:dP AEGIS pair is stronger than the dC:dG pair, and comparable to that of the diso^{Me}C:disoG pair.[8] In addition, the dZ:dP AEGIS pair effectively discriminates against mismatches in short duplex DNA.[8] Furthermore, the dZ:dP base pair is readily accepted by many native DNA polymerases.[10] All these features make the dZ:dP pair a greater AEGIS pair for future applications in developing advanced diagnostic tools.

To implement and benchmark this approach, four AEGIS-containing MBs and a control (ZPMB-1, ZPMB-2C, ZPMB-2S, ZPMB-3 and Control MB, Table 1) were prepared by chemical synthesis from the appropriate phosphoramidites.[8] Each MB sequence in the series was analogous to the previous, but had an additional dC:dG pair in its stem replaced by a dZ:dP pair. Standard conditions for automated DNA synthesis were used, except that extended coupling times were used for the dZ and dP nucleoside phosphoramidites (6 min) and the 6-fluorescein phosphoramidite (15 min). For deprotection, the modified MBs were first treated with DBU (1 M in acetonitrile, 12 h, room temperature) to remove the protecting group on dZ (p-nitrophenethyl (NPE)), then, incubated in NH₃(aq) (29%) at 55 °C overnight. All molecular beacons

were purified by preparative reverse-phase HPLC and their purities were checked by analytical ion-exchange HPLC.

The T_m for opening of each beacon was first determined and compared. All MBs modified with dZ and dP showed higher T_ms than the control MB (T_m = 57.8 °C). This was consistent with our previous work showing that dZ:dP base pair is stronger than dC:dG pair in duplex DNA.[8] Melting temperatures indicate that the ZPMB-2S (T_m = 68.2 °C) and ZPMB-3 (T_m = 67.8 °C) form more stable hairpins than ZPMB-1 (T_m = 60.5 °C) and ZPMB-2C (T_m = 61.3 °C), suggesting that increased substitution of dC:dG pairs by dZ:dP pairs creates stepwise more stable hairpins. By comparing the T_ms of ZPMB-3 with ZPMB-2S, and the T_ms of ZPMB-2C with ZPMB-1, we found that the stability added by substituting two consecutive AEGIS pairs (the dinucleotide dPdZ paired with the dinucleotide dZdP) approximately equals that added by substituting just one dP:dZ pair. This was unexpected given previous studies that showed that the thermal stability of two consecutive pairs in duplex DNA (dPdP:dZdZ) is more than additive, showing a degree of context dependence with AEGIS components comparable to that found in standard DNA.

The signal enhancement observed after the AEGIS-containing MBs hydridize to their complementary fully matched target (binding in the loop), and the thermal denaturation profiles of MBs bound to its loop targets were then measured. Upon hybridization with their targets, all the AEGIS-containing MBs show comparable signal enhancements with the control MB (lacking AEGIS-components). Further, both modified and control MBs give consistently reasonable thermal denaturation profiles.

To show that MB stems containing dZ and dP resisted stem invasion and thereby reduced false-positive signals, MBs were challenged with various synthetic oligonucleotides with different lengths that had six nucleotides complementary to the stem of the MB, and a remaining sequence matched the loop region (Table 1). Fluorescence emission was recorded after separately incubating these with both modified and control MBs for 0.5 h. For the control MB, incubation with 10-fold excess of the 6mer and 8mer DNA target (Table 1) failed to open the hairpin structure (Fig. 3a, Control MB, lane 2 and 3). This is consistent with the design strategy; the intramolecular-binding

Table 1 Sequences of MBs and targets used in this study

Name	Sequence
Control MB	5'-FAM-<u>CCTAGC</u>TCTAAATCACTATGGTCGCG<u>CTAGG</u>-Dabcyl-3'
ZPMB-1	5'-FAM-<u>CCTAGZ</u>TCTAAATCACTATGGTCGC<u>PCTAGG</u>-Dabcyl-3'
ZPMB-2C [a]	5'-FAM-<u>CCTAPZ</u>TCTAAATCACTATGGTCGC<u>PZTAGG</u>-Dabcyl-3'
ZPMB-2S [b]	5'-FAM-<u>CZTAPC</u>TCTAAATCACTATGGTCGC<u>GZTAPG</u>-Dabcyl-3'
ZPMB-3	5'-FAM-<u>CZTAPZ</u>TCTAAATCACTATGGTCGC<u>PZTAPG</u>-Dabcyl-3'
Loop Target	3'-AGATTTAGTGATACCAGCG-5'
6mer Target	3'-<u>CGATCC</u>-5'
8mer Target	3'-CG<u>CGATCC</u>-5'
10mer Target	3'-AGC<u>GCGATCC</u>-5'
12mer Target	3'-CCAGC<u>GCGATCC</u>-5'
ZP-10mer Target	3'-<u>GPATZ</u>GAGAT-5'
ZP-12mer Target	3'-<u>GPATZ</u>GAGATTT-5'

[a] 2C Indicates that the MB has two consecutively modified nucleobases in the stem. [b] 2S Indicates that the MB has two separately modified nucleobases in the stem. Underlined regions are stems and their targets. The presence of Z in the synthetic oligonucleotides was confirmed by UV absorbance at 390 nm.
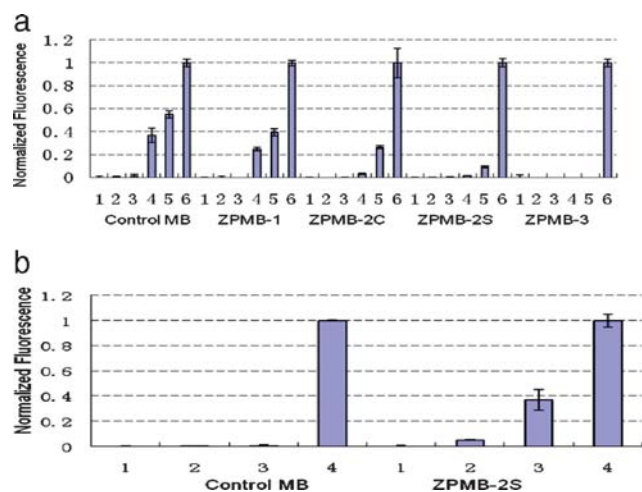
Fig. 3 Comparison of selectivity of control MB and modified MBs. A final concentration of 50 nM MB was incubated with 500 nM of each target for 0.5 h and the fluorescence signal was measured. The experiment was repeated three times and the average value was calculated. (a) MBs (lane 1, normalized as zero); MBs incubate with the targets (6mer, 8mer, 10mer, and 12mer target, respectively, lane 2, 3, 4, and 5); MBs incubate with the loop targets (lane 6, normalized as one). (b) MBs (lane 1, normalized as zero); MBs incubate with the targets (ZP-10mer and ZP-12mer target, lane 2 and 3); MBs incubate with the loop targets (lane 4, normalized as one).

MB with a higher T_m, ZPMB-2S, generates a better signal-to-noise ratio than the one with lower T_m, ZPMB-2C.

To show that the modified MBs eliminate stem invasion mainly because of the orthogonality of AEGIS pairing, and not because of the higher T_m associated with dZ:dP pairs in the stem, two sequences were designed to have complementarity to one stem in the AEGIS-containing MBs (Table 1). Specifically, the ZP-10mer target and ZP-12mer target, replaced one dC and one dG with one dZ and one dP, respectively. As shown in Fig. 3(b), the control MB is not opened by either the ZP-10mer or the ZP-12mer target (lane 2 or 3). However, the ZPMB-2S beacon is partially opened by ZP-12mer target (land 3), even though the T_m of the modified MB is about 10 °C higher than that of the control MB.

In conclusion, we have shown that the use of AEGIS-components in the stems of MB can eliminate unwanted stem invasion by adventitious DNA. The selectivity of the modified MBs arises mainly from the orthogonality of dZ:dP pair over natural base pairs. The stability of the hairpin stem increase stepwise 3.2 °C in average, after an additional dC:dG pair in the stem replaced by a dZ:dP pair. This should further mitigate any background by alternative beacon conformations involving stem–loop interactions.

constant between the stem sequences is far greater than the intermolecular interaction between one arm of the stem and the 6mer and 8mer DNA targets. On the other hand, in the presence of the 10mer and 12mer targets, one of the stem and part of loop of the control MB hybridizes with the target, disrupting the hairpin structure to give about 38% and 56% (assuming a linear fluorescence response) of opened MB, respectively (Fig. 3(a), Control MB, lane 4 and 5).

In contrast, the AEGIS-containing MBs cannot be opened by any of the stem targets tested (Fig. 3(a), ZPMB-3). Indeed, results of the experiments shown in Fig. 3(a) suggested that the ability of an AEGIS-containing MB to avoid the stem invasion issue is directly related to both the number of AEGIS components in the stem and the melting temperature of the stem. In general, three AEGIS substitutions are better than two, and two are better than one. As for ZPMB-3 and ZPMB-2S, the thermal stability of each stem is roughly the same, however, the ability of ZPMB-3 to resist stem invasion is superior to that of ZPMB-2S. The selectivity of hybridization is presumably due to the orthogonality of AEGIS pairing, which is directly related to the number of AEGIS-components in the stem. On the other hand, with two substitutions, the modified

Notes and references

1 S. Tyagi and F. R. Kramer, *Nat. Biotechnol.*, 1996, **14**, 303.
2 (a) J. Li, X. H. Fang, S. M. Schuster and W. H. Tan, *Angew. Chem., Int. Ed.*, 2000, **39**, 1049; (b) W. H. Tan, X. H. Fang, J. Li and X. J. Liu, *Chem.–Eur. J.*, 2000, **6**, 1107.
3 (a) M. M. Mhlanga, D. Y. Vargas, C. W. Fung, F. R. Kramer and S. Tyagi, *Nucleic Acids Res.*, 2005, **33**, 1902; (b) C. Molenaar, S. A. Marras, J. C. M. Slats, J. C. Truffert, M. Lemaitre, A. K. Raap, R. W. Dirks and H. J. Tanke, *Nucleic Acids Res.*, 2001, **29**, e89.
4 (a) O. Seitz, *Angew. Chem., Int. Ed.*, 2000, **39**, 3249; (b) C. W. Xi, M. Balberg, S. A. Boppart and L. Raskin, *Appl. Environ. Microbiol.*, 2003, **69**, 5673; (c) L. Wang, C. Y. J. Yang, C. D. Medley, S. A. Benner and W. H. Tan, *J. Am. Chem. Soc.*, 2005, **127**, 15664; (d) C. Crey-Desbiolles, D. R. Ahn and C. J. Leumann, *Nucleic Acids Res.*, 2005, **33**, e77.
5 P. J. Santangelo, B. Nix, A. Tsourkas and G. Bao, *Nucleic Acids Res.*, 2004, **32**, e57.
6 Y. Kim, C. Y. J. Yang and W. H. Tan, *Nucleic Acids Res.*, 2007, **35**, 7279.
7 S. A. Benner, *Acc. Chem. Res.*, 2004, **37**, 784.
8 Z. Yang, D. Hutter, P. Sheng, A. M. Sismour and S. A. Benner, *Nucleic Acids Res.*, 2006, **34**, 6095.
9 T. Elbeik, J. Surtihadi, M. Destree, J. Gorlin, M. Holodniy, S. A. Jortani, K. Kuramoto, V. Ng, R. Valdes, A. Valsamakis and N. A. Terrault, *J. Clin. Microbiol.*, 2004, **42**, 563.
10 Z. Yang, A. M. Sismour, P. Sheng, N. L. Puskar and S. A. Benner, *Nucleic Acids Res.*, 2007, **35**, 4238.