Single-Nucleotide Polymorphism

DOI: 10.1002/anie.201004475

Real-Time SNP Analysis in Secondary-Structure-Folded Nucleic Acids**

*Jeffrey Grimes, Yulia V. Gerasimova, and Dmitry M. Kolpashchikov**

Hybridization of two complementary nucleic acid strands is extensively used in the analysis of specific DNA/RNA sequences in real-time PCR, DNA microarrays, and techniques for RNA monitoring in living cells. The design of the hybridization probes is based on A–T and G–C complementarity and may seem straightforward. However, single-stranded DNA and RNA analytes often form stable secondary structures under assay conditions. The analysis of such folded analytes is often complicated because a region of interest may be involved in intramolecular hybridization and become inaccessible for hybridization with a probe. This complication severely limits sensitivity and creates and insurmountable obstacle for the detection of single nucleotide differences between two analytes.[1] Herein we demonstrate an approach that allows analysis of single-nucleotide polymorphisms (SNPs) in folded analytes in real time at room temperature.

SNPs are the most abundant class of genetic variations in humans, accounting for 80–90% of the difference between the genomes of two individuals.[2] There are over 3 million validated SNPs, although the actual number is believed to be in the range of 10 million. SNP analysis is important in population-based genetic risk assessment, molecular diagnostics, pharmaceutical drug development, linkage analysis, and identity testing in forensic applications.[3]

An important advantage in nucleic acid analysis came with the introduction of real-time hybridization probes,[4] which allow instant fluorescent signal readout without the need to separate the probe–analyte hybrid from the excess of unbound probes. A state-of-the-art tool for real-time nucleic acid analysis is the molecular beacon (MB) probe,[4d,5] which is a fluorophore- and quencher-conjugated DNA hairpin (Figure 1a). Binding to a specific analyte separates the fluorophore from the quencher, thus enabling a fluorescence increase. Herein we demonstrate that MB probe is inefficient in detection of stem-loop folded analytes and suggest an alternative approach for the analysis of such problematic analytes in real time.

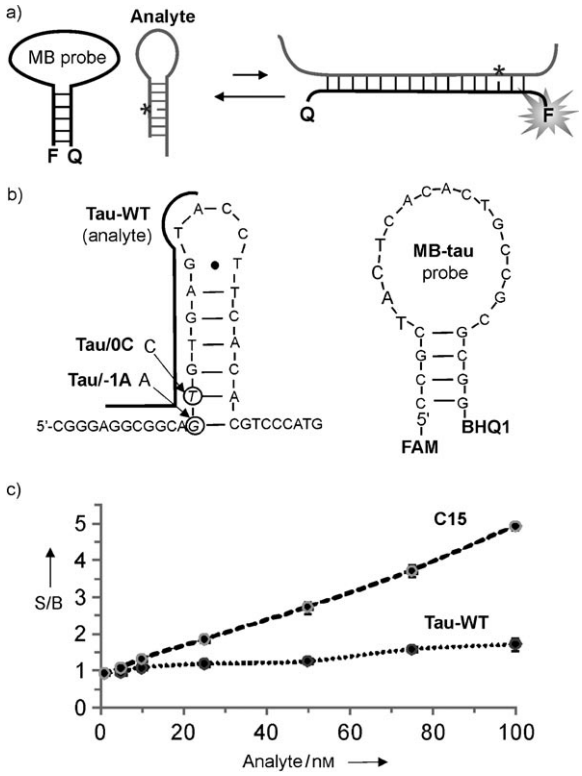

Figure 1. Molecular beacon probe for the analysis of Tau-WT analyte. a) The hybridization of MB probe to a hairpin-folded analyte. The SNP site is indicated by an asterisk. The equilibrium is shifted toward the dissociated state. b) The primary and secondary structures of Tau-WT analyte and MB-tau. Two naturally SNPs are indicated and designated as Tau/0C and Tau/-1A. The nucleotides complementary to MB-tau are indicated by the bar. c) The fluorescent signal-to-background (S/B) ratios of MB-tau in the presence of different concentrations of Tau-WT (•••••) and a hairpin structure-free oligonucleotide C15 (GCG GCA GTG TGA GTA), which is complementary to the loop fragment of MB-tau (-----). Average values of three independent measurements with one standard deviation as error bars are presented.

As a model analyte we chose a fragment of the gene that encodes tau protein (Tau-WT, Figure 1b). It was shown earlier that SNPs in this sequence cause genetic predisposition for Alzheimer's disease.[6] It is important that the disease is linked with SNPs located within a specific stem structure, such as Tau/0C and Tau/-1A (Figure 1b). A single mutation that reduces stem stability leads to the alternative mRNA splicing, which eventually causes the production of the mutant protein capable of forming nerve tissue-distracting filaments.

An MB probe containing a loop segment complementary to Tau-WT analyte was designed (MB-tau; Figure 1b). MB-tau was unable to generate a significant fluorescence increase

[*] J. Grimes, Dr. Y. V. Gerasimova, Dr. D. M. Kolpashchikov
Chemistry Department, University of Central Florida
4000 Central Blvd, Orlando, FL 32816 (USA)
Fax: (+1) 407-823-2252
E-mail: dkolpash@mail.ucf.edu

[**] Support from the UCF Office of Research and Commercialization, College of Science, Chemistry Department at UCF and NHGRI R21 HG004060 is greatly appreciated. SNP = single-nucleotide polymorphism.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201004475.

in the presence of Tau-WT analyte (Figure 1c, dotted curve). A maximum S/B ratio of about 1.5 in the presence of 100 nM Tau-WT was detected. In contrast, addition of an equivalent amount of C15 (an oligonucleotide complementary to MB-tau), which did not form secondary structure, generated significant concentration dependent fluorescence increase (Figure 1c, dashed curve).

This result demonstrates the low efficiency of the MB probe to report the presence of secondary-structure-folded analytes by fluorescence. Indeed, the MB probe and the analyte exist preferably in the dissociated state at the concentrations used. The dissociated forms were stabilized by the formation of the four-base-pair stem in MB-tau probe and the six-base-pair stem plus the wobble G-T base pair in the analyte (Figure 1b). This explanation was supported by the observations that neither annealing nor increased incubation times resulted in higher fluorescence (data not shown), thus indicating weak interaction of MB-tau probe with the analyte at the concentrations used (ca. 100 nM). A significant fluorescence increase was observed only when the analyte concentration was increased up to about 1000 nM. It should be noted that normally MB probes detect circa 1–5 nM targets that do not form secondary structures.[5] This severe reduction in sensitivity of the MB approach is undesirable for the majority of practical applications. For the analytes that form strong secondary structures, the reduced hybridization efficiency and the low sensitivity of linear hybridization probes have been reported previously.[1]

To enable SNP analysis of secondary structure-complicated analytes, we suggest the following general approach. The analyte is incubated with a universal MB (UMB) probe and a set of two target-specific adaptor oligonucleotides, strand f and strand m (Figure 2). Strands m and f contain fragments complementary to the analyte (analyte-binding arms) and fragments complementary to the MB probe (MB-binding arms). In the absence of specific nucleic acid analyte, the MB-binding arms of strands m and f only weakly interact with the MB probe because each arm hybridizes to a relatively short (9 nucleotide) fragment of UMB. In the presence of analyte, the long analyte-binding arm of strand f hybridizes to the fragment adjacent to the SNP site and thus unwinds the analyte secondary structure. The resultant complex contains the SNP site in a single-stranded region that is accessible for analysis. Subsequent hybridization of strand m and the MB probe results in the formation of a quadripartite complex, which contains MB probe in highly fluorescent open conformation. Importantly, strand m contains a short analyte-binding arm that forms a stable hybrid only with the fully complementary analyte. The presence of a single base mismatch in the case of mutation will destabilize the quadripartite complex, resulting in its decomposition into MB probe, strand m, and strand f–analyte duplex. In this proof-of-concept study we demonstrate that this approach indeed allows SNP analysis in tau-protein-encoding nucleic acid fragment.

UMB probe, strand m, and strand f were incubated in the presence of different concentrations of Tau-WT analyte for 15 min at room temperature followed by fluorescent measurements (Figure 3, solid curve). The concentrations of

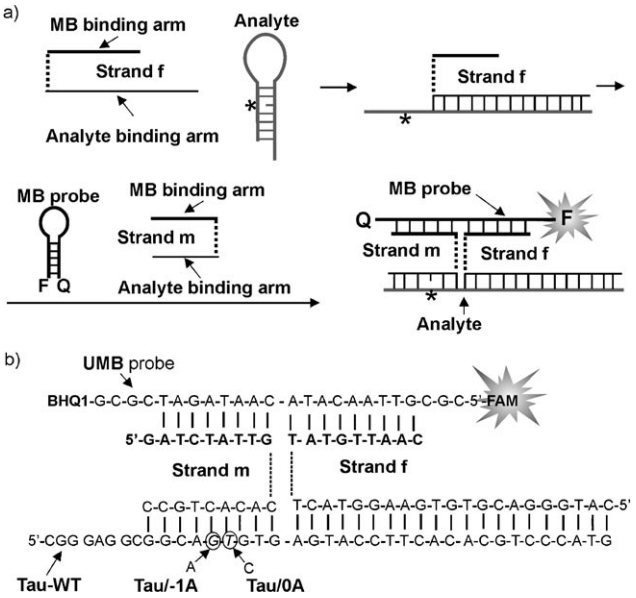

Figure 2. The tricomponent MB-based probe for SNP analysis in secondary structure-forming analytes. a) Probe design. Strands f and m contain analyte-binding arms and MB-binding arms. The SNP site in the analyte is indicated by an asterisk. b) Tripartite probe for the analysis of Tau protein-encoding sequence. The primary and secondary strictures of the universal MB (UMB) probe and the quadripartite complex with the analyte are shown. The mutated nucleotide positions are indicated.

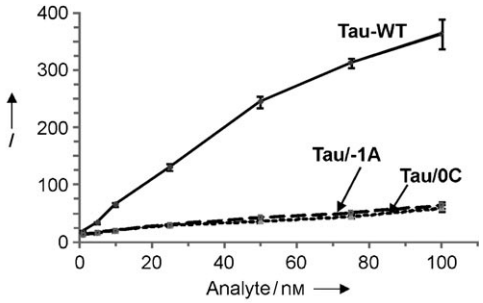

Figure 3. Fluorescence intensities of the tricomponent MB-based probe in the presence of different concentrations of the fully complementary analyte Tau-WT (——) or one of the single-base mismatched analytes Tau/0C (-----) and Tau/-1A (– – –).

adaptors strands m and f were pre-optimized to produce the highest possible signal-to-background ratio (S/B). Therefore, at the concentrations of strands used, a significant background in the absence of the analyte was avoided, whilst allowing the highest possible analyte-dependent fluorescence increase. The fluorescent signal increased in a concentration-dependent manner and reached a S/B of 9 at the analyte concentration of 100 nM. The signal was significantly above the background even in the presence of 5 nM analyte. This sensitivity corresponds to the sensitivity of the conventional MB probes.[5] When a single-base-mismatched analyte Tau/0C or Tau/-1A was added, the fluorescence of the samples was close to background levels (Figure 3, dashed lines). These results show excellent selectivity of the MB-based tripartite

probe, which is in agreement with the earlier findings of high hybridization specificity of binary and multicomponent probes.[7] These data support the concept for the analysis of secondary-structure-folded nucleic acids suggested in Figure 2. Importantly, not only the probe detects the analyte in real time, but it is also capable of accurate SNP genotyping at ambient temperatures.

To demonstrate the general applicability of the approach, T13 analyte, an oligonucleotide that forms a stem–loop structure with 13 nucleotide-long stem was designed (Supporting Information, Figure S1, Table S1). The design of the adaptor strands was straightforward: the analyte-binding arms of the new strands sequences were made to be complementary to the new analyte, whilst the MB-binding arms remained the same to allow optimal interaction with UMB. It was found that anti-T13 probe efficiently detected T13 with high signal-to-background ratio (Supporting Information, Figure S2). Like the probe described above, the anti-T13 construct efficiently discriminated single-base-substituted oligonucleotide targets against T13 analyte (Supporting Information, Figure S2). These data suggest that the proposed approach is applicable for analytes of arbitrary nucleic acid sequence, and even long stem fragments do not necessarily prevent accurate SNP genotyping under mild hybridization conditions.

It is of practical interest to analyze nucleic acids that are involved in hybridization. For example, Frank-Kamenetskii and co-workers have been developing an elegant approach that takes advantage of PNA oligomers to open double-stranded DNA and make it available for hybridization with MB probe.[8] The study presented herein demonstrates that self-assembling multicomponent probes can be designed to analyze secondary-structure-folded nucleic acids. In these probes, each constructing element is specialized to serve a distinct function. Strand f forms a stable complex with the analyte, which unwinds its secondary (and perhaps tertiary) structure. Strand m forms a short sensitive to a mismatch hybrid with the SNP site and, therefore, is responsible for the probe specificity. The MB-probe reports the formation of the quadripartite complex. It is not surprising that splitting the functional roles between the three specialized components results in a superior tool for nucleic acid analysis. It is important that this approach is cost-efficient in the case of multiple targets analysis. Indeed, the sequence of the universal MB probe is independent on the analyte sequence and, therefore, it can be synthesized in bulk amounts and used efficiently for the analysis of any DNA or RNA. At the same time, analyte-specific strands m and f are relatively inexpensive oligonucleotides that can be used without HPLC purification. To sum up, this study demonstrates advantages of a smart muticomponent probe in which each component serves distinct function. This design may find a number of applications in nucleic acid detection techniques of the future.

Experimental Section

All of the oligonucleotides were custom-made by Integrated DNA Technologies, Inc (Coralville, IA). For the fluorescence assay with MB-tau probe, MB-tau and an analyte were mixed in a buffer containing 2 mM $MgCl_2$, 50 mM Tris-HCl, pH 7.4, at a final concentration of 100 nM for the MB-tau probe and 1–100 nM for the analytes. For the fluorescence assay with a tricomponent probe, strand m, strand f, UMB, and an analyte were mixed in a buffer containing 20 mM $MgCl_2$, 50 mM Tris-HCl, pH 7.4, with strands m and f, and UMB being at 100 nM and the analytes being at 1–100 nM. Fluorescence spectra were recorded on a Perkin–Elmer (San Jose, CA) LS-55 luminescence spectrometer with a Hamamatsu xenon lamp (excitation at 485 nm; emission 517 nm) after 15 min of incubation at room temperature (22 °C). The data of three independent measurements are presented with an error margin of one standard deviation.

Received: July 21, 2010
Revised: September 9, 2010
Published online: October 20, 2010

Keywords: biotransformations · genotyping · molecular beacons · real-time detection · tau proteins

[1] a) J. C. Tan, J. J. Patel, A. Tan, J. C. Blain, T. J. Albert, N. F. Lobo, M. T. Ferdig, *Genomics* **2009**, *93*, 543–550; b) S. A. Kushon, J. P. Jordan, J. L. Seifert, H. Nielsen, P. E. Nielsen, B. A. Armitage, *J. Am. Chem. Soc.* **2001**, *123*, 10805–10813; c) B. A. Armitage, *Drug Discovery Today* **2003**, *8*, 222–228.

[2] D. Altshuler, L. D. Brooks, A. Chakravarti, F. S. Collins, M. J. Daly, P. Donnelly, *Nature* **2005**, *437*, 1299–1320.

[3] a) G. C. Johnson, J. A. Todd, *Curr. Opin. Genet. Dev.* **2000**, *10*, 330–334; b) N. J. Schork, D. Fallin, J. S. Lanchbury, *Clin. Genet.* **2000**, *4*, 250–264; c) A. L. Beaudet, J. W. Belmont, *Annu. Rev. Med.* **2008**, *59*, 113–129; d) C. Güzey, O. Spigset, *Curr. Top. Med. Chem.* **2004**, *4*, 1411–1421; e) N. J. Risch, *Nature* **2000**, *405*, 847–856; f) S. B. Gabriel, S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, *Science* **2002**, *296*, 2225–2229; g) B. Budowle, A. van Daal, *Biotechniques* **2008**, *44*, 603–608; h) D. M. Altshuler, R. A. Gibbs, and the international HapMap 3 consortium, *Nature* **2010**, *467*, 52–58.

[4] a) R. A. Cardullo, S. Agrawal, C. Flores, P. C. Zamechnik, D. E. Wolf, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8790–8794; b) L. E. Morrison, T. C. Halder, L. M. Stols, *Anal. Biochem.* **1989**, *183*, 231–244; c) K. Ebata, M. Masuko, H. Ohtani, M. Kashiwasake-Jibu, *Photochem. Photobiol.* **1995**, *62*, 836–839; d) S. Tyagi, F. R. Kramer, *Nat. Biotechnol.* **1996**, *14*, 303–308; e) A. P. Silverman, E. T. Kool, *Adv. Clin. Chem. Trends Biotechnol.* **2005**, *23*, 225–230.

[5] a) S. A. Marras, S. Tyagi, F. R. Kramer, *Clin. Chim. Acta* **2006**, *363*, 48–60; b) K. Wang, Z. Tang, C. J. Yang, Y. Kim, X. Fang, W. Li, Y. Wu, C. D. Medley, Z. Cao, J. Li, P. Colon, H. Lin, W. Tan, *Angew. Chem.* **2009**, *121*, 870–885; *Angew. Chem. Int. Ed.* **2009**, *48*, 856–870; c) Y. Li, X. Zhou, D. Ye, *Biochem. Biophys. Res. Commun.* **2008**, *373*, 457–461; d) N. Venkatesan, Y. J. Seo, B. H. Kim, *Chem. Soc. Rev.* **2008**, *37*, 648–663.

[6] a) L. Varani, M. Hasegawa, M. G. Spillantini, M. J. Smith, J. R. Murrell, B. Ghetti, A. Klug, M. Goedert, G. Varani, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8229–8234; b) M. Goedert, R. Jakes, *Biochim. Biophys. Acta Gen. Subj.* **2005**, *1721*, 240–250.

[7] a) J. L. Mergny, A. S. Boutorine, T. Garestier, F. Belloc, M. Rougée, N. V. Bulychev, A. A. Koshkin, J. Bourson, A. V. Lebedev, B. Valeur, N. T. Thuong, C. Helene, *Nucleic Acids Res.* **1994**, *22*, 920–928; b) A. Oser, G. Valet, *Angew. Chem.* **1990**, *102*, 1197–1200; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1167–1169; c) K. Ebata, M. Masuko, H. Ohtani, M. Kashiwasake-Jibu, *Photochem. Photobiol* **1995**, *62*, 836–839; d) A. P. Silverman, E. T. Kool, *Chem. Rev.* **2006**, *106*, 3775–3789; e) Y. Kitamura, T. Ihara, Y. Tsujimura, Y. Osawa, D. Sasahara, M. Yamamoto, K. Okada, M. Tazaki, A. Jyo, *J. Inorg. Biochem.* **2008**, *102*, 1921–1931; f) E. V. Bichenkova, H. E. Savage, A. R. Sardarian, K. T.

Douglas, *Biochem. Biophys. Res. Commun.* **2005**, *332*, 956 – 964; g) D. M. Kolpashchikov, *J. Am. Chem. Soc.* **2005**, *127*, 12442 – 12443; h) D. M. Kolpashchikov, *J. Am. Chem. Soc.* **2006**, *128*, 10625 – 10628; i) D. M. Kolpashchikov, *ChemBioChem* **2007**, *8*, 2039 – 2042; j) D. M. Kolpashchikov, *J. Am. Chem. Soc.* **2008**, *130*, 2934 – 2935; k) Y. V. Gerasimova, E. Cornett, D. M. Kolpashchikov, *ChemBioChem* **2010**, *11*, 811 – 817; l) D. M. Kolpashchikov, *Chem. Rev.* **2010**, *110*, 4709 – 4723; m) Y. V. Gerasimova, A. Hayson, J. Ballantyne, D. M. Kolpashchikov, *ChemBioChem* **2010**, *11*, 1762 – 1768.

[8] a) H. Kuhn, V. V. Demidov, J. M. Coull, M. J. Fiandaca, B. D. Gildea, M. D. Frank-Kamenetskii, *J. Am. Chem. Soc.* **2002**, *124*, 1097 – 1103; b) I. V. Smolina, V. V. Demidov, V. A. Soldatenkov, S. G. Chasovskikh, M. D. Frank-Kamenetskii, *Nucleic Acids Res.* **2005**, 33, e146; c) I. V. Smolina, H. Kuhn, C. Lee, M. D. Frank-Kamenetskii, *Bioorg. Med. Chem.* **2008**, *16*, 84 – 93.

