Published in final edited form as:

Angew Chem Int Ed Engl. 2010 November 15; 49(47): 8950–8953. doi:10.1002/anie.201004475.

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

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

Chemistry Department, University of Central Florida, 4000 Central Blvd, Orlando, FL 32816 (USA), Fax: (+1) 407-823-2252, dkolpash@mail.ucf.edu

Keywords

hybridization probe; SNP genotyping; molecular beacon probe; Tau protein; real-time detection

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 the 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 since 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 an insurmountable obstacle for the detection of single nucleotide differences between two analytes.[1] This study demonstrates an approach that allows analysis of single nucleotide polymorphisms (SNPs) in folded analytes in real time at room temperature.

SNPs represent 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 probe-analyte hybrid from the excess of unbound probes. A state-of-the-art tool for real time nucleic acid analysis is molecular beacon (MB) probe.[4d,5] MB probe is a fluorophore- and a quencher-conjugated DNA hairpin (Figure 1a). Binding to a specific analyte separates fluorophore from the quencher, thus enabling fluorescence increase. In this study 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.

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

**Support from the UCF Office of Research and Commercialization, College of Science, Chemistry Department at UCF and NHGRI R21 HG004060 is greatly appreciated

Correspondence to: Dmitry M. Kolpashchikov.

Supporting information for this article is available on the WWW under <http://www.angewandte.org> or from the authors.

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

This result demonstrates the low efficiency of MB probe to fluorescently report the presence of secondary structure-folded analytes. Indeed, 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 time resulted in higher fluorescence (data not shown), thus indicating weak interaction of **MB-tau** probe with the analyte at the concentrations used (~100 nM). Significant fluorescent increase was observed only when the analyte concentration was increased up to ~ 1000 nM. It should be noted that normally MB probes detect ~ 1–5 nM targets that do not form secondary structures.[5] This severe reduction in sensitivity of 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 earlier.[1]

In order 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 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 accessible for analysis. Subsequent hybridization of **strand m** and 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 which 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, dotted curve). The concentrations of adaptors **strands m** and **f** were pre-optimized to produce the highest possible signal-to-background ratio (S/B). Therefore, at the strands' concentrations used, significant background in the absence of the analyte was avoided, while allowing the highest possible analyte-dependent fluorescence

increase. 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 the background (Figure 3, dashed lines). These results prove 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.

In order 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 (Figure S1 and Table S1). The design of the adaptor strands was straightforward: the analyte-binding arms of the new strands sequences were made complementary to the new analyte, while 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 (Figure S2). Like the probe described above, the anti-**T13** construct efficiently discriminated single base substituted oligonucleotide targets against **T13** analyte (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's group has 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 here 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. 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, which can be used without HPLC purification. To sum up, this study demonstrates advantages of a smart multicomponent 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 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 MgCl₂, 50 mM Tris-HCl, pH 7.4, at a final concentration of 100 nM for **MB-tau** probe and 1–100 nM for the analytes. For the fluorescence assay with a tri-component probe, **strand m**, **strand f**, **UMB**, and an analyte were mixed in a buffer containing 20 mM MgCl₂, 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. a) Tan JC, Patel JJ, Tan A, Blain JC, Albert TJ, Lobo NF, Ferdig MT. *Genomics*. 2009; 93:543–550. [PubMed: 19285129] b) Kushon SA, Jordan JP, Seifert JL, Nielsen H, Nielsen PE, Armitage BA. *J Am Chem Soc*. 2001; 123:10805–10813. [PubMed: 11686681] c) Armitage BA. *Drug Discov Today*. 2003; 8:222–228. [PubMed: 12634014]
2. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. *Nature*. 2005; 437:1299–1320. [PubMed: 16255080]
3. a) Johnson GC, Todd JA. *Curr Opin Genet Dev*. 2000; 10:330–334. [PubMed: 10826983] b) Schork NJ, Fallin D, Lanchbury JS. *Clin Genet*. 2000; 4:250–264. [PubMed: 11076050] c) Beaudet AL, Belmont JW. *Ann Rev Med*. 2008; 59:113–129. [PubMed: 17961075] d) Güzey C, Spigset O. *Curr Top Med Chem*. 2004; 4:1411–1421. [PubMed: 15379654] e) Risch NJ. *Nature*. 2000; 405:847–856. [PubMed: 10866211] f) Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A. *Science*. 2002; 296:2225–2229. [PubMed: 12029063] g) Budowle B, van Daal A. *Biotechniques*. 2008; 44:603–608. [PubMed: 18474034] h) Altshuler DM, Gibbs RA. the international HapMap 3 consortium. *Nature*. 2010; 467:52–58. [PubMed: 20811451]
4. a) Cardullo RA, Agrawal S, Flores C, Zamechnik PC, Wolf DE. *Proc Natl Acad Sci USA*. 1988; 85:8790–8794. [PubMed: 3194390] b) Morrisson LE, Halder TC, Stols LM. *Anal Biochem*. 1989; 183:231–244. [PubMed: 2624314] c) Ebata K, Masuko M, Ohtani H, Kashiwasake-Jibu M. *Photochem Photobiol*. 1995; 62:836–839. [PubMed: 8570721] d) Tyagi S, Kramer FR. *Nat Biotechnol*. 1996; 14:303–308. [PubMed: 9630890] e) Silverman AP, Kool ET. *Adv Clin Chem Trends Biotechnol*. 2005; 23:225–230.
5. a) Marras SA, Tyagi S, Kramer FR. *Clin Chim Acta*. 2006; 363:48–60. [PubMed: 16111667] b) Wang K, Tang Z, Yang CJ, Kim Y, Fang X, Li W, Wu Y, Medley CD, Cao Z, Li J, Colon P, Lin H, Tan W. *Angew Chem*. 2009; 121:870–885. *Angew Chem Int Ed Engl*. 2009; 48:856–870. [PubMed: 19065690] c) Li Y, Zhou X, Ye D. *Biochem Biophys Res Commun*. 2008; 373:457–461. [PubMed: 18489905] d) Venkatesan N, Seo YJ, Kim BH. *Chem Soc Rev*. 2008; 37:648–663. [PubMed: 18362974] e) Kushon SA, Jordan JP. *Am Chem Soc*. 2001; 123:10805–10813.
6. a) Varani L, Hasegawa M, Spillantini MG, Smith MJ, Murrell JR, Ghetti B, Klug A, Goedert M, Varani G. *Proc Natl Acad Sci USA*. 1999; 96:8229–8234. [PubMed: 10393977] b) Goedert M, Jakes R. *Biochim Biophys Acta*. 2005:240–250. [PubMed: 15615642]
7. a) Cardullo RA, Agrawal S, Flores C, Zamechnik PC, Wolf DE. *Proc Natl Acad Sci USA*. 1988; 85:8790–8794. [PubMed: 3194390] b) Mergny JL, Bourtoune AS, Garestier T, Belloc F, Rougée M, Bulychev NV, Koshkin AA, Bourson J, Lebedev AV, Valeur B, Thuong NT, Helene C. *Nucleic Acids Res*. 1994; 22:920–928. [PubMed: 8152922] c) Oser A, Valet G. *Angew Chem Int Ed Engl*. 1990; 29:1167–1169. d) Ebata K, Masuko M, Ohtani H, Kashiwasake-Jibu M. *Photochem Photobiol*. 1995; 62:836–839. [PubMed: 8570721] e) Silverman AP, Kool ET. *Chem Rev*. 2006; 106:3775–3789. [PubMed: 16967920] f) Kitamura Y, Ihara T, Tsujimura Y, Osawa Y, Sasahara D, Yamamoto M, Okada K, Tazaki M, Jyo A. *J Inorg Biochem*. 2008; 102:1921–1931. [PubMed: 18707760] g) Bichenkova EV, Savage HE, Sardarian AR, Douglas KT. *Biochem Biophys Res Commun*. 2005; 332:956–964. [PubMed: 15913549] h) Kolpashchikov DM. *J Am Chem Soc*. 2005; 127:12442–12443. [PubMed: 16144363] i) Kolpashchikov DM. *J Am Chem Soc*. 2006; 128:10625–10628. [PubMed: 16895431] j) Kolpashchikov DM. *ChemBioChem*. 2007; 8:2039–2042. [PubMed: 17924377] k) Kolpashchikov DM. *J Am Chem Soc*. 2008; 130:2934–2935. [PubMed: 18281995] l) Gerasimova YV, Cornett E, Kolpashchikov DM. *ChemBioChem*. 2010; 11:811–817. [PubMed: 20301161] m) Kolpashchikov DM. *Chem Rev*. 2010; 110:4709–4723. [PubMed: 20583806] n)

- Gerasimova YV, Hayson A, Ballantyne J, Kolpashchikov DM. *Chembiochem*. 2010; 11:1762–1768. [PubMed: 20665615]
8. a) Kuhn H, Demidov VV, Coull JM, Fiandaca MJ, Gildea BD, Frank-Kamenetskii MD. *JACS*. 2002; 124:1097–1103. b) Smolina IV, Demidov VV, Soldatenkov VA, Chasovskikh SG, Frank-Kamenetskii MD. *Nucleic Acids Res*. 2005; 33:e146. [PubMed: 16204449] c) Smolina IV, Kuhn H, Lee C, Frank-Kamenetskii MD. *Bioorg Med Chem*. 2008; 16:84–93. [PubMed: 17512202]

Figure 1.

Molecular beacon probe for the analysis of Tau-WT analyte. a) General scheme of hybridization of MB probe to a hairpin-folded analyte. 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** (dotted curve) and a hairpin structure-free oligonucleotide **C15** (GCG GCA GTG TGA GTA), which is complementary to the loop fragment of **MB-tau** (dashed curve). Average values of three independent measurements with one standard deviation as error bars are presented.

Figure 2.

The tri-component MB-based probe for SNP analysis in secondary structure-forming analytes. a) Probe design. **Stand**s **f** and **m** contain analyte-binding arms and MB-binding arms. 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.

Fluorescent intensities of the tri-component MB-based probe in the presence of different concentrations of the fully complementary analyte **Tau-WT** (dotted curve) or one of the single-base mismatched analytes **Tau/0C** and **Tau/-1A** (dashed curves).