Supporting information for the paper:

Environmentally sensitive molecular probes reveal mutations and epigenetic 5-methyl cytosine in human oncogenes

Maria Taskova ,[a] Maria Carla Barducci,[a] and Kira Astakhova,[a]

[a.] *Nucleic Acid Center, Department of Physics, Chemistry and Pharmacy, Campusvej 55, Odense M 5230, Denmark*
E-mail: ias@sdu.dk

Table of Contents

1. General

Reagents were obtained from commercial suppliers and used as received. Fluorescent azides and TBTA for click chemistry were obtained from Lumiprobe GmbH. Stock solutions for click chemistry were prepared as described.[1] Click reactions were performed in 1.5 mL Eppendorf tubes or 1 mL reactor tubes under argon using vigorous stirring in Emrys Creator (Personal Chemistry) in the latter case.

Precursor probes used for click chemistry and for *in vitro* analyses were synthesised at NAC SDU, Odense, as described in the literature.[2] Precursors **C2** and **C3** were purified on IE HPLC; other strands did not require purification (purity > 90%, see Table S2). DNA/RNA targets were purchased from IDT Integrated DNA Technologies. Cell line DNA (HMC, LS411N, GP2D and RCM1) was generously provided by Dr. Hanlee Ji, Stanford University, USA.

2. Probes: preparation and analyses

We clicked six **X**-labelled oligonucleotides with three fluorescent dyes for a total of 18 product probes. The precursor strands contain none LNA (**C1**), or three LNA building blocks (**C2-C6**), and contain two commercially available scaffolds with an alkyne group for the click reaction (Chart 1).

Sequences of precursors obtained from pubmed genbank public database (Gene ID: 673 and 3845) are listed below.

C1 (*BRAF* V600E m) 5'-CGAGA**X**TTCTCTGTAGC**X**AGA-3'
C2 (*BRAF* V600E m)5'-CGAGA**X**TTC^{L}T^{L}C^{L}TGTAGC**X**AGA-3'
C3 (*BRAF* V600E w) 5'-CGAGA**X**TTC^{L}A^{L}C^{L}TGTAGC**X**AGA-3'

C4 (KRAS G12D m) 5'-GTX GGA GCT +G+A+T GGC GXA GGC-3'
C5 (KRAS G13D m) 5'-GTX GGA GCT GGT +G+A+C GXA GGC-3'
C6 (KRAS G12/13D w) 5'-GTX GGA +GCT G+GT +GGC GXA GGC-3'

[1] Web: http://www.lumiprobe.com/protocols/click-chemistry-dna-labeling
[2] Okholm et al. RCS Adv. 2014

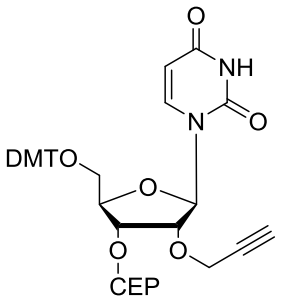


Chart 1 alkyne modified scaffold (**X**)

Table S1. Codes for the click chemistry products prepared in this work.

Starting oligonucleotide	Attached dye:		
	Perylene	6FAM	5JOE
C1	**1**	**2**	**3**
C2	**4**	**5**	**6**
C3	**7**	**8**	**9**
C4	**10**	**11**	**12**
C5	**13**	**14**	**15**
C6	**16**	**17**	**18**

Click chemistry of the oligonucleotides with 5JOE and 6FAM dyes was conducted at room temperature. Microwave assisted reaction were run to click perylene dye.

Starting oligonucleotide (20 nmol in MQ water) was mixed with 20 μL of buffer triethylammonium acetate (TEAA) buffer (1M in MQ water), dimethyl sulfoxide DMSO, 24 μL of dye (10 mM in DMSO), MQ water. At this point the solution was flushed carefully with argon. Then 10 μL of Cu(II)-TBTA (10mM in MQ water-DMSO 1:1) was added and the solution was flushed with argon a second time. 4 μL of fresh prepared ascorbic acid solution (50 mM in MQ water) were added in the end. The final reaction mixture was calculated to have a volume of 200 μL, with 60% of DMSO and 40% of MQ water for all the reactions.

The mixtures were closed tightly and vortexed. Perylene samples were put in a microwave reactor for 30 minutes at 60°C (the initial power was between 40 and 50 W for every reaction, it decreased to 12-15 W once the selected temperature was reached). 5JOE and 6FAM samples underwent a hot start, i.e. mixed at 400 rpm at 75°C for 10 minutes. The resulting mixtures were stored at room

temperature overnight. Afterwards perylene samples were exposed to additional microwave treatment for 10 minutes at 60°C, whereas 6FAM and 5JOE samples were additionally shacked by hand. All the samples were left at room temperature overnight. Next, all the samples were precipitated with cold acetone, followed by redissolving in MQ water and filtration through NAP^{TM}-10 columns (Sephadex^{TM} G-25 DNA Grade, GE Healthcare) following manufacturer's protocol. The concentration of the products was calculated by measuring the OD_{260}.[2]

Probe analyses

All the samples were analysed by MALDI-MS and IE HPLC (Table S2, Figs. S1-S6). For HPLC we used two programmes, PS and PO (PS: instrument - Merck Hitachi LaChrom equipped with Dionex DNAPac Pa-100 column (250 mm × 4 mm) at room temperature, elution - starting with an isocratic hold of A- and C-buffers for 2 min followed by a linear gradient to 60% B-buffer over 28 min at a flow rate of 1.0 mL/min (A-buffer: MQ water; B-buffer: 1M $LiClO_4$, C-buffer: 250 mM Tris-Cl, pH 8.0); PO: instrument - Merck Hitachi LaChrom equipped with Dionex DNAPac Pa-100 column (250 mm × 4 mm) at 60 °C, elution - gradient: 0–60% B in A within 45 min (A: 25 mM Tris·HCl (pH 8.0), 6 M urea; B: 25 mM Tris·HCl (pH 8.0), 0.5 M $NaClO_4$, 6 M urea), flow rate of 1.0 mL/min; UV detection at 254 nm for both PO and PS programmes). PS programme was used for the *BRAF* oligonucleotides that did not show any secondary structure formation in most cases. For perylene-labelled *BRAF* probes **1** and **7**, the two observed peaks were separated by PS IE HPLC. However they gave similar mass in MALDI MS (Figure S6). Therefore the major portion of the product probes was used without purification.

KRAS precursors **C4**-**C6** showed multiple peaks in PS IE HPLC and single mass in the MALDI MS. We changed the HPLC conditions to the PO programme that was run at 60 °C. This programme is useful for the analysis of oligonucleotides forming secondary structures.[3] Indeed, we observed single peaks for the precursors **C4**-**C6** under the PS IE HPLC conditions (Figure S2). Products **10**-**18** were analysed by PS IE HPLC as well (Figures S3-S4).

[3] Flür S, Micura R. *Methods (San Diego, Calif)*. 2016;107:79-88.

MALDI-TOF mass-spectrometry analysis was performed using a MALDI-LIFT system on the Ultraflex II TOF/TOF instrument from Bruker and using HPA-matrix (10 mg 3-hydroxypicolinic acid, 50 mM ammonium citrate in 70% aqueous acetonitrile). Being compared to the calculated mass, standard error for the obtained values is m/z = 2.95, which is 0.04 % of the average mass for the analysed oligonucleotides (m/z = 7352; Table S2). This confirms high accuracy of the MALDI MS analysis for the obtained probes.

Table S2. Characterization of oligonucleotides by IE HPLC and MALDI MS.[a]

Precursor/ Probe nr.	Probe composition: Precursor + Dye	MS calcd.	MS found	HPLC ret time, min	Purity by HPLC, %
C1	-	6514	6511	13.25	95
C2	-	6626	6629	12.42	88(90)[b]
C3	-	6635	6630	12.08	82(91)[b]
C4	-	6720	6722	***10.61***	***95***
C5	-	6734	6730	***9.76***	***95***
C6	-	6736	6738	***9.90***	***92***
1	**C1** + 2x Per	7184	7189	14.55; 15.02	93
2	**C1** + 2x 6FAM	7442	7440	16.32	91
3	**C1** + 2x 5JOE	7690	7695	17.35	92
4	**C2** + 2x Per	7304	7301	15.49	90
5	**C2** + 2x 6FAM	7552	7554	14.79	92
6	**C2** + 2x 5JOE	7809	7806	16.71	88
7	**C3** + 2x Per	7307	7307	14.31; 14.77	90
8	**C3** + 2x 6FAM	7542	7534	15.07	96
9	**C3** + 2x 5JOE	7800	7794	16.18	90
10	**C4** + 2x Per	7390	7393	***18.82***	***91***
11	**C4** + 2x 6FAM	7650	7651	***16.78***	***89***
12	**C4** + 2x 5JOE	7910	7911	***18.76***	***94***
13	**C5** + 2x Per	7390	7392	***16.70***	***96***
14	**C5** + 2x 6FAM	7650	7651	***15.86***	***93***

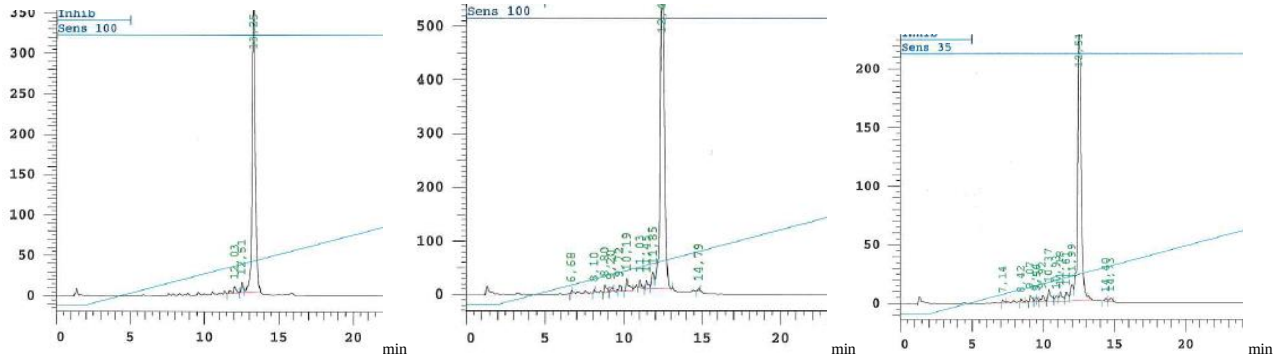
15	**C5** + 2x 5JOE	7910	7913	***17.89***	***93***
16	**C6** + 2x Per	7390	7387	***18.96***	***91***
17	**C6** + 2x 6FAM	7650	7647	***18.09***	***90***
18	**C6** + 2x 5JOE	7910	7911	***19.44***	***93***

[a] **C1-C6** are listed on page S3. HPLC retention times and purity for PS and PO programmes are shown in regular style and bold italic. For **C2-C3**, the conditions were similar for the purification and analysis. Products probes are analysed with no purification, right after the click reaction work up.

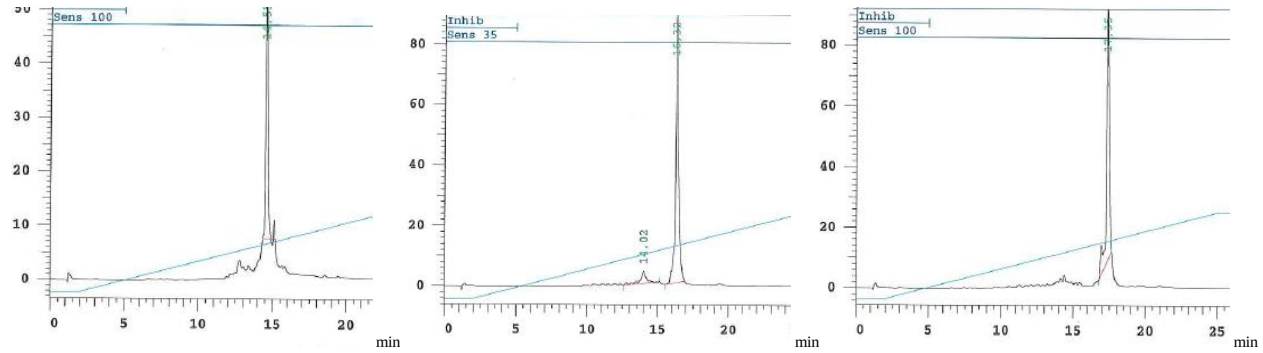
[b] purity assessment before(after) the IE HPLC purification.

Figure S1. IE HPLC spectra of *BRAF* precursors C1-C3 and product probes, PS programme.

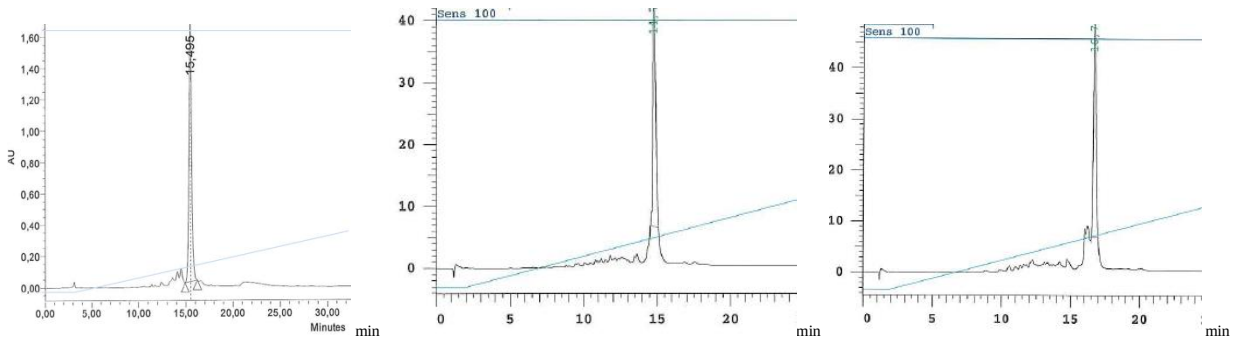
C1 **C2[a]** **C3[a]**



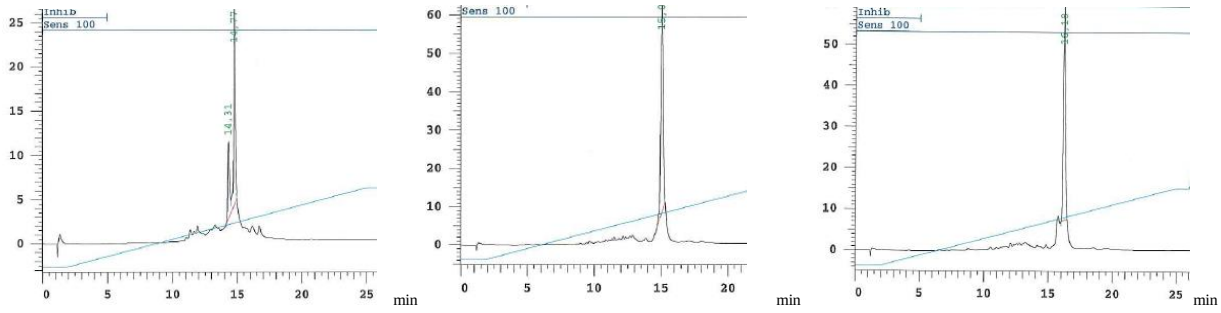
1 **2** **3**



4 **5** **6**



7 **8** **9**



[a] HPLC spectra for the precursors **C2-C3** are given before the purification.

Figure S2. IE HPLC spectra of *KRAS* precursors C4-C6, PO programme

PO IE HPLC: C4 **C5** **C6**

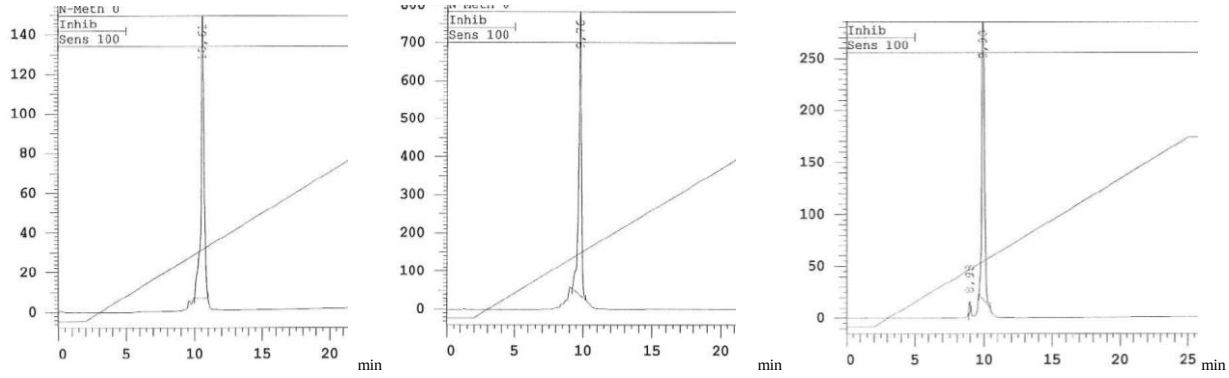
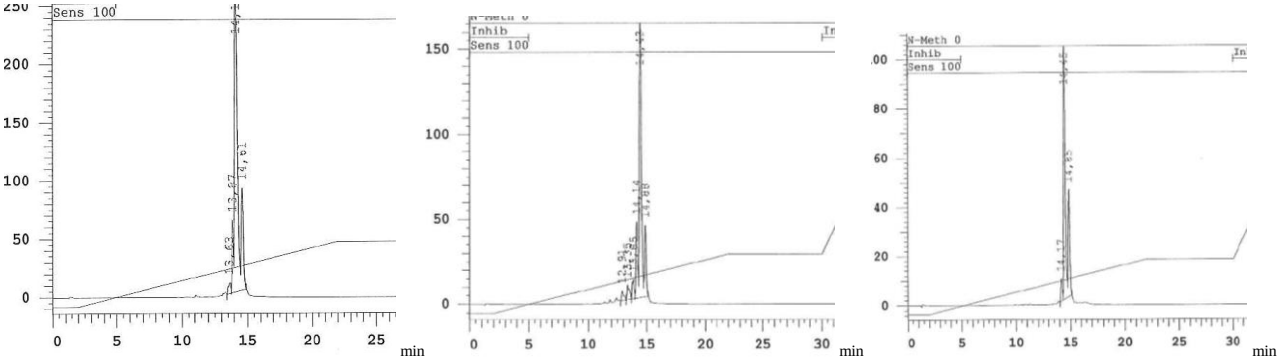


Figure S3. IE HPLC spectra of *KRAS* product probes labelled with perylene, PS and PO programmes

PS programme: 10 **13** **16**



PO programme: 10 **13** **16**

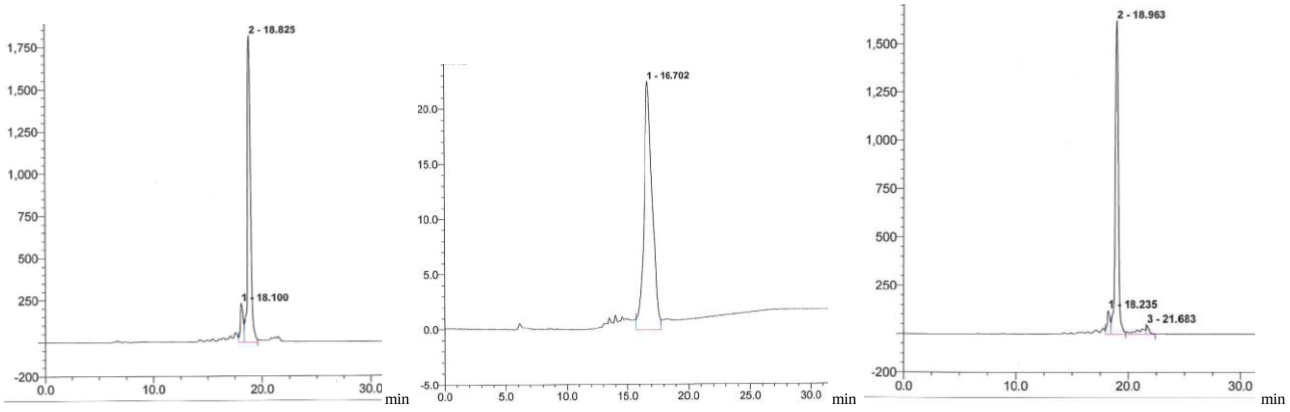


Figure S4. IE HPLC spectra of *KRAS* product probes labelled with 6FAM and 5JOE, PO programme

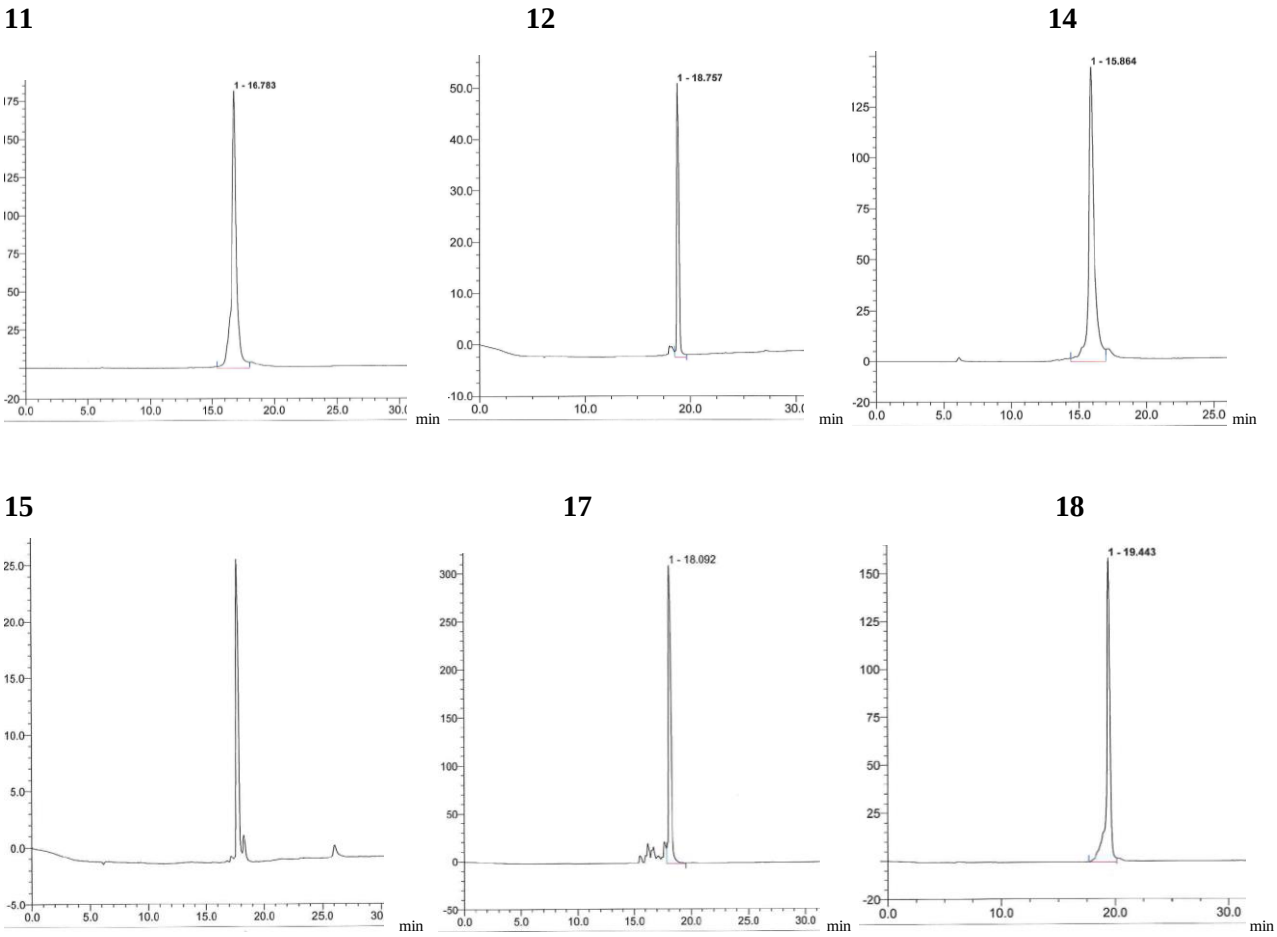
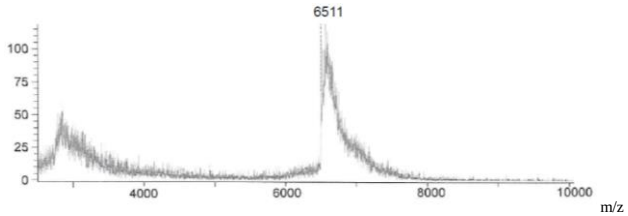
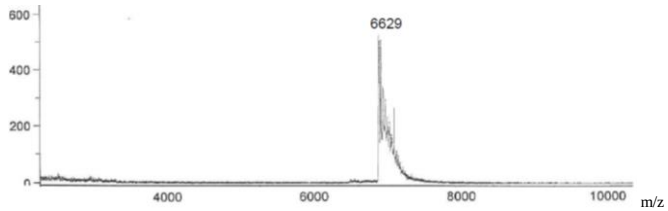


Figure S5. MALDI MS spectra of C1-C3 precursors and product probes. 1’ and 7’ correspond to the second peaks in the HPLC profilesof probes 1 and 7.

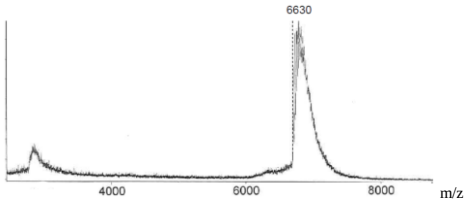
C1



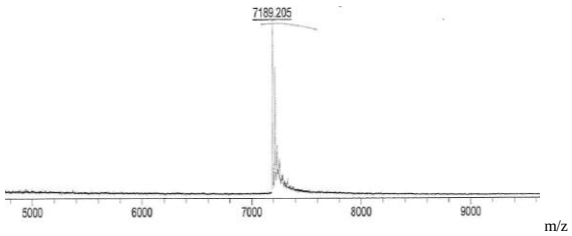
C2



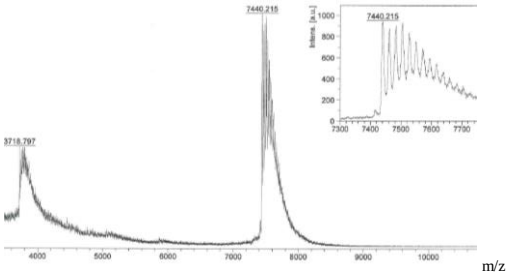
C3



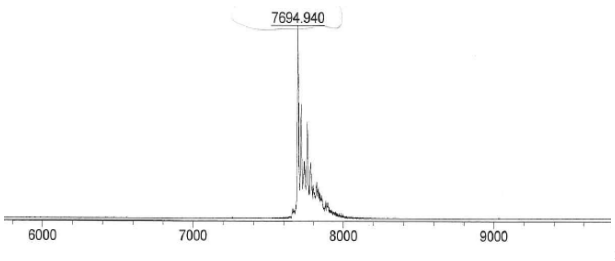
1



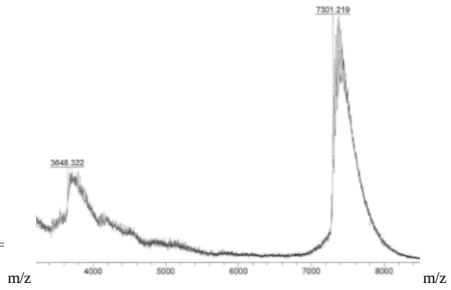
2



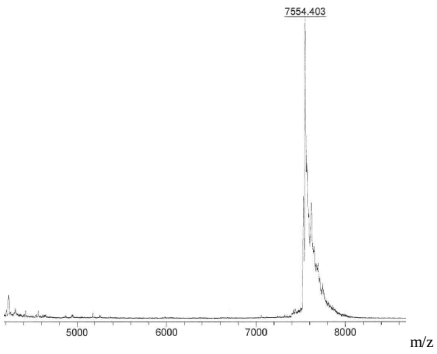
3



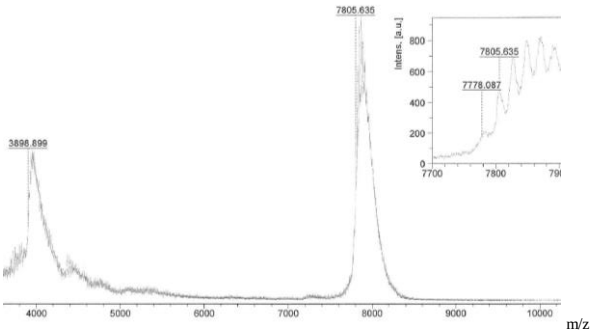
4



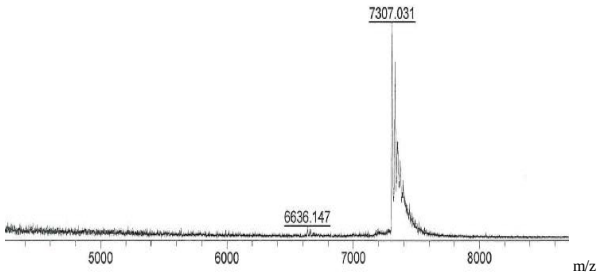
5


7554.403
5000
6000
7000
8000
m/z

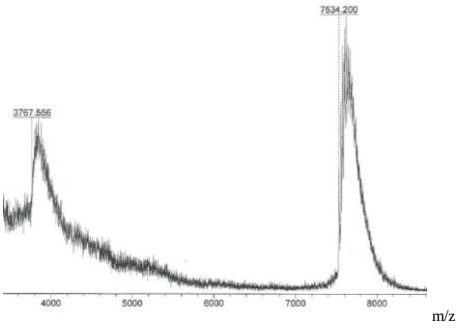
6


7805.636
3898.899
Intens. [a.u.]
800
600
400
200
0
7805.635
7778.087
7700
7800
4000
5000
6000
7000
8000
9000
10000
m/z

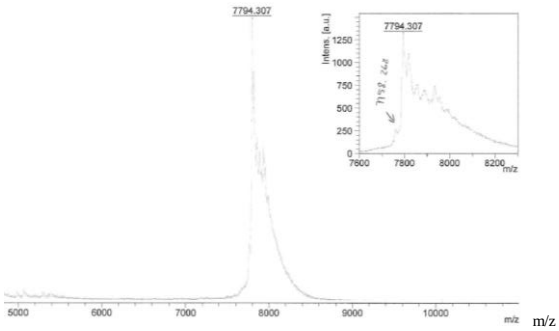
7


7307.031
6636.147
5000
6000
7000
8000
m/z

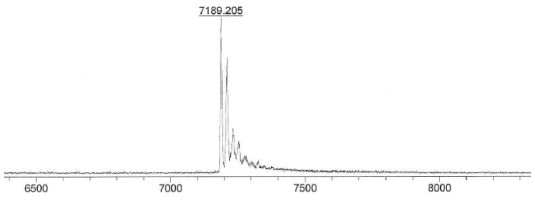
8


7534.200
3767.556
4000
5000
6000
7000
8000
m/z

9


7794.307
Intens. [a.u.]
1250
1000
750
500
250
0
7794.307
7600
7800
8000
8200
m/z
5000
6000
7000
8000
9000
10000
m/z

1’


7189.205
6500
7000
7500
8000

7’

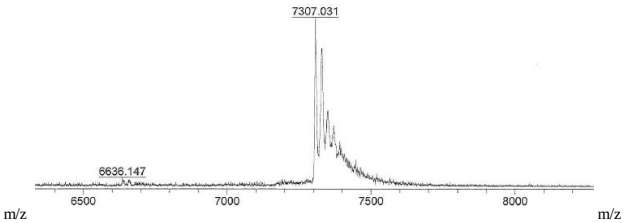
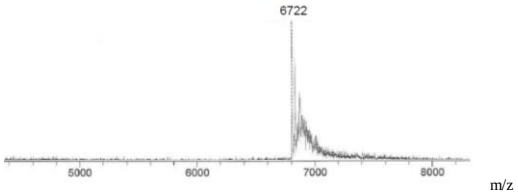
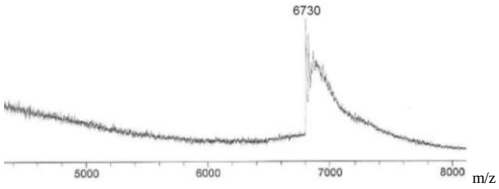

7307.031
6636.147
6500
7000
7500
8000
m/z

Figure S6. MALDI MS spectra of C4-C6 precursors and product probes

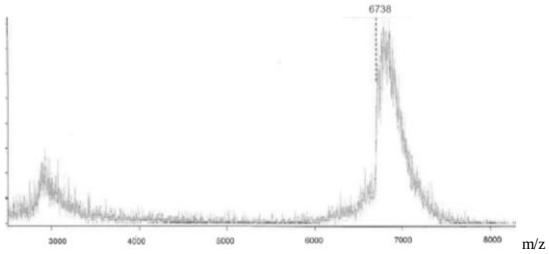
C4



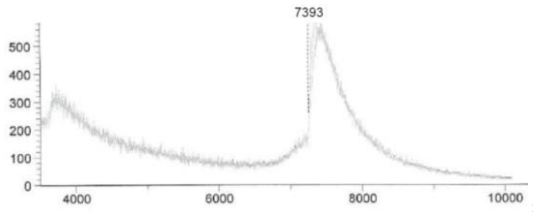
C5



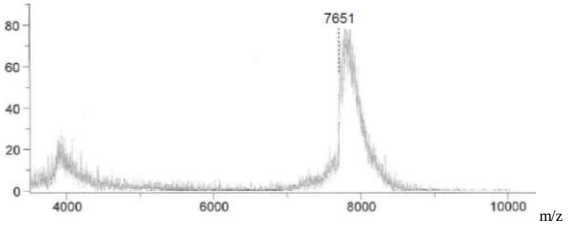
C6



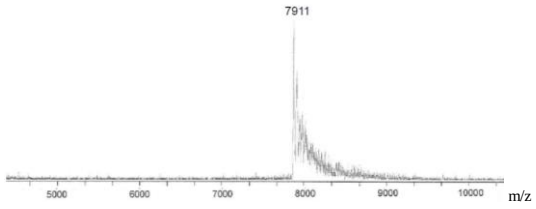
10



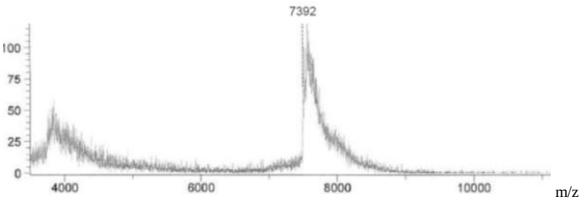
11



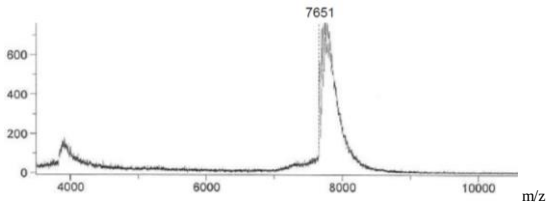
12



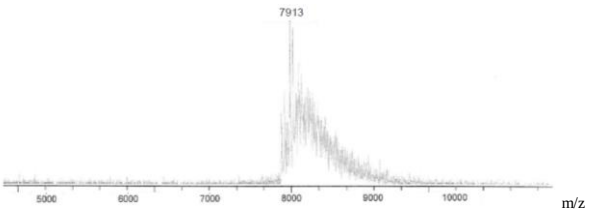
13



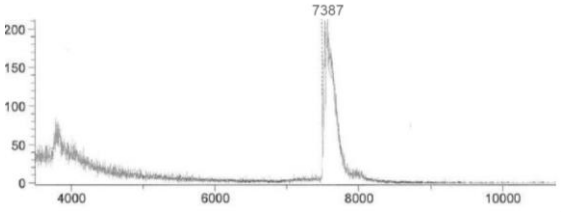
14



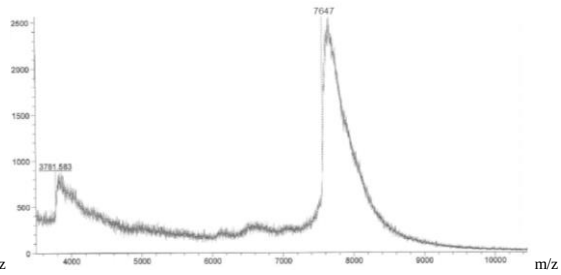
15



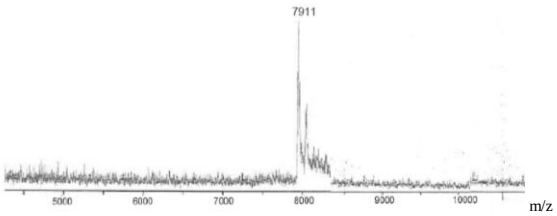
16


7387
200
150
100
50
0
4000
6000
8000
10000
m/z

17


7647
3781.583
2500
2000
1500
1000
500
0
4000
5000
6000
7000
8000
9000
10000
m/z

18


7911
5000
6000
7000
8000
9000
10000
m/z

3. T_m and CD studies

3.1. Probe annealing

All the samples were annealed with complementary and mismatched DNA/RNA strands listed below. Mutation position in each target is underlined.

BRAF targets:
V600E D1: 5'-TCT AGC TAC AGA GAA ATC TCG-3' (wild type)
V600E D2: 5'-TCT AGC TAC AG<u>T</u> GAA ATC TCG-3' (mutant)

V600E R1: 5'-rUrCrU rArGrC rUrArC rA rGrA rGrArA rArUrC rUrCrG -3' (wild type)
V600E R2: 5'-rUrCrU rArGrC rUrArC rArG<u>r</u>U rGrArA rArUrC rUrCrG -3' (mutant)

KRAS targets:
G12D D1: 5'-GCCTACG<u>C</u>CATCAGCTCCAAC-3' (mutant)
G13D D1: 5'-GCCTACGTCA<u>C</u>CAGCTCCAAC-3' (mutant)
G12D/G13D D2: 5'-GCCTACGCCACCAGCTCCAAC-3' (wild type)

G12D R1: 5'-rGrCrCrUrArCrGrCr<u>C</u>rArUrCrArGrCrUrCrCrArArC-3' (mutant)
G13D R1: 5'-rGrCrCrUrArCrGrUrCrAr<u>C</u>rCrArGrCrUrCrCrArArC-3' (mutant)
G12D/G13D R2: 5'-rGrCrCrUrArCrGrCrCrArCrCrArGrCrUrCrCrArArC-3' (wild type)

For annealing, the two strands were mixed in equimolar ratio in 1xPBS, pH 7.2, vortexed, kept at 95°C for 10 min followed by cooling to room temperature over 4 hr.

3.2. T_m measurements

Studies were carried out using DU®800 spectrophotometer, Beckman Coulter. All ds samples were analyzed twice in a concentration 0.5 μM in 1xPBS buffer, pH 7.2.

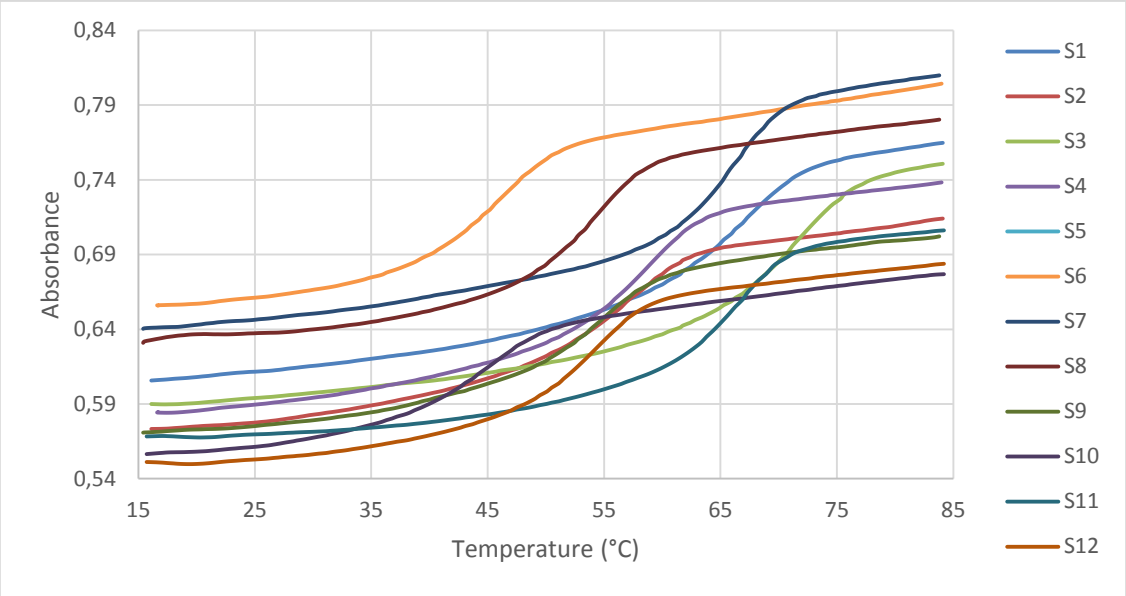


Figure S7. Thermal denaturation curves of duplexes containing *BRAF* specific fluorescent probes hybridized with matching and mismatching DNA and RNA. Legend's abbreviations refer to Table S3.

Table S3 Codes of T_m samples hybridized with matching (M) and mismatching (MS) *BRAF* DNA and RNA.[a]

Probe number/composition	Target DNA/RNA:			
	D1 (M)	D2 (MS)	R1 (M)	R2 (MS)
1/C1+ Per	S1	S2	S3	S4
5/C2+ 6FAM	S5	S6	S7	S8
9/C3 + 5JOE	S9	S10	S11	S12

[a] *BRAF* targets: V600E D1: 5'-TCT AGC TAC AGA GAA ATC TCG-3' (wild type); V600E D2: 5'-TCT AGC TAC AG<u>T</u> GAA ATC TCG-3' (mutant); V600E R1: 5'-rUrCrU rArGrC rUrArC rA rGrA rGrArA rArUrC rUrCrG -3' (wild type); V600E R2: 5'-rUrCrU rArGrC rUrArC rArGr<u>U</u> rGrArA rArUrC rUrCrG -3' (mutant). Mutation position is underlined. See page S3 for the sequences of precursors C1-C3.

Table S4. Yields and T_m data for DNA recognition by *KRAS* specific probes[a]

Probe no.	Probe sequence, 5'→3'	Target	Yield (%)	T_m/Δ T_m,[b] °C vs. DNA
10	GT**M**$^{\mathbf{1}}$GGAGCT^{L}G^{L}A^{L}TGGCG**M**$^{\mathbf{1}}$AGGC	*KRAS12m*	85	66.0/-12.0
11	GT**M**$^{\mathbf{2}}$GGAGCT^{L}G^{L}A^{L}TGGCG**M**$^{\mathbf{2}}$AGGC	*KRAS12m*	88	60.0/-8.0
13	GT**M**$^{\mathbf{1}}$GGAGCT^{L}G^{L}A^{L}TGGCG**M**$^{\mathbf{1}}$AGGC	*KRAS13m*	82	64.5/-11.0
14	GT**M**$^{\mathbf{2}}$GGAGCT^{L}G^{L}A^{L}TGGCG**M**$^{\mathbf{2}}$AGGC	*KRAS13m*	87	57.0/-10.0
16	GT**M**$^{\mathbf{1}}$GGA^{L}GCTG^{L}GT^{L}GGCG**M**$^{\mathbf{1}}$AGGC	*KRASw*	81	68.0/-8.5
17	GT**M**$^{\mathbf{2}}$GGA^{L}GCTG^{L}GT^{L}GGCG**M**$^{\mathbf{2}}$AGGC	*KRASw*	84	64.0/-11.0

[a] For the chemical structures of **M**$^{\mathbf{1}}$-**M**$^{\mathbf{3}}$ and **X**, see Fig. 1 and Chart 1. *KRAS* targets: G12D D1: 5'-GCCTACG<u>C</u>CATCAGCTCCAAC-3' (mutant), G13D D1: 5'-GCCTACGTCA<u>C</u>CAGCTCCAAC-3' (mutant), G12D/G13D D2: 5'-GCCTACGCCACCAGCTCCAAC-3' (wild type). Mutation position is underlined. Δ T_m is calculated for each probe as the difference between fully matched and mismatched probe:target duplexes.

Table S5. T_m data for RNA recognition by *BRAF* specific probes[a]

Probe no.	Probe sequence, 5'→3'	Target	T_m/Δ T_m,[b] °C vs. RNA
1	CGAGA**M**$^{\mathbf{1}}$TTCTCTGTAGC**M**$^{\mathbf{1}}$AGA	*BRAFm*	59.0/-9.0
2	CGAGA**M**$^{\mathbf{2}}$TTCTCTGTAGC**M**$^{\mathbf{2}}$AGA	*BRAFm*	57.0/-8.0
3	CGAGA**M**$^{\mathbf{3}}$TTCTCTGTAGC**M**$^{\mathbf{3}}$AGA	*BRAFm*	57.0/-8.0
4	CGAGA**M**$^{\mathbf{1}}$TTC^{L}T^{L}C^{L}TGTAGC**M**$^{\mathbf{1}}$AGA	*BRAFm*	72.0/-13.0
5	CGAGA**M**$^{\mathbf{2}}$TTC^{L}T^{L}C^{L}TGTAGC**M**$^{\mathbf{2}}$AGA	*BRAFm*	66.0/-12.0
6	CGAGA**M**$^{\mathbf{3}}$TTC^{L}T^{L}C^{L}TGTAGC**M**$^{\mathbf{3}}$AGA	*BRAFm*	67.0/-12.0
7	CGAGA**M**$^{\mathbf{1}}$TTC^{L}A^{L}C^{L}TGTAGC**M**$^{\mathbf{1}}$AGA	*BRAFw*	70.0/-8.0
8	CGAGA**M**$^{\mathbf{2}}$TTC^{L}A^{L}C^{L}TGTAGC**M**$^{\mathbf{2}}$AGA	*BRAFw*	64.0/-8.5
9	CGAGA**M**$^{\mathbf{3}}$TTC^{L}A^{L}C^{L}TGTAGC**M**$^{\mathbf{3}}$AGA	*BRAFw*	64.0/-8.0
C1	CGAGA**X**TTCTCTGTAGC**X**AGA	*BRAFm*	59.9/-7.3
C2	CGAGA**X**TTC^{L}T^{L}C^{L}TGTAGC**X**AGA	*BRAFm*	70.1/-11.6
C3	CGAGA**X**TTC^{L}A^{L}C^{L}TGTAGC**X**AGA	*BRAFw*	65.9/-9.9

[a] For the chemical structures of **M**$^{\mathbf{1}}$-**M**$^{\mathbf{3}}$ and **X**, see Fig. 1 and Chart 1. *BRAF* targets: V600E R1: 5'-rUrCrU rArGrC rUrArC rA rGrA rGrArA rArUrC rUrCrG -3' (wild type), V600E R2: 5'-rUrCrU rArGrC rUrArC rArGrU rGrArA rArUrC rUrCrG -3' (mutant). Mutation position is underlined. Δ T_m is calculated for each probe as the difference between fully matched and mismatched probe:target duplexes.

Table S6. T_m data for RNA target recognition by *KRAS* specific probes[a]

Probe no.	Probe sequence, 5'→3'	Target	T_m/Δ T_m,[b] °C vs. RNA
10	GT**M^1** GGAGCT^{L}G^{L}A^{L}TGGCG**M^1**A GGC	*KRAS12m*	70.0/-11.0
11	GT**M^2** GGAGCT^{L}G^{L}A^{L}TGGCG**M^2**A GGC	*KRAS12m*	63.5/-12.0
12	GT**M^3** GGAGCT^{L}G^{L}A^{L}TGGCG**M^3**A GGC	*KRAS12m*	64.0/-10.5
13	GT**M^1** GGAGCT^{L}G^{L}A^{L}TGGCG**M^1**A GGC	*KRAS13m*	71.0/-11.5
14	GT**M^2** GGAGCT^{L}G^{L}A^{L}TGGCG**M^2**A GGC	*KRAS13m*	61.5/-9.5
15	GT**M^3** GGAGCT^{L}G^{L}A^{L}TGGCG**M^3**A GGC	*KRAS13m*	62.0/-10.0
16	GT**M^1**GGA^{L}GCTG^{L}GT^{L}GGCG**M^1**AGGC	*KRASw*	68.0/-11.0
17	GT**M^2**GGA^{L}GCTG^{L}GT^{L}GGCG**M^2**AGGC	*KRASw*	65.5/-11.5
18	GT**M^3**GGA^{L}GCTG^{L}GT^{L}GGCG**M^3**AGGC	*KRASw*	61.0/-13.5
C3	GT**X**GGAGCT^{L}G^{L}A^{L}TGGCG**X**AGGC	*KRAS12m*	77.0/-8.0
C4	GT**X** GGAGCT^{L}G^{L}A^{L}TGGCG**X**A GGC	*KRAS13m*	78.0/-5.5
C5	GT**X**GGA^{L}GCTG^{L}GT^{L}GGCG**X**AGGC	*KRASw*	72.0/-6.5

[a] For the chemical structures of **M^1-M^3** and **X**, see Fig. 1 and Chart 1. *KRAS* targets: G12D R1: 5'-rGrCrCrUrArCrGrCrCrArUrCrArGrCrUrCrCrArArC-3' (mutant), G13D R1: 5'-rGrCrCrUrArCrGrUrCrArCrCrArGrCrUrCrCrArArC-3' (mutant), G12D/G13D R2: 5'-rGrCrCrUrArCrGrCrCrCrArCrCrArGrCrUrCrCrArArC-3' (wild type). Mutation position is underlined. Δ T_m is calculated for each probe as the difference between fully matched and mismatched probe:target duplexes.

Table S7. Difference in T_m due to the presence of LNA in the strand of the fluorescent probe targeting *BRAF*. M=matching. MS=mismatching.[a]

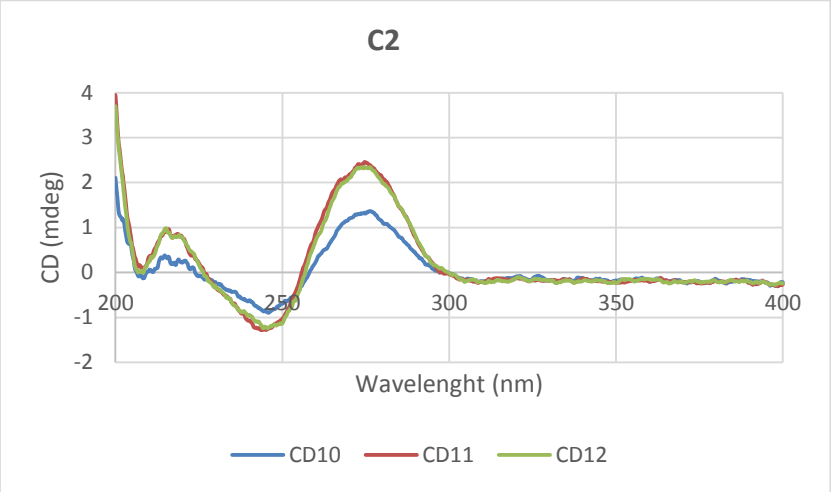
Target type (M/MS)	Calculation of LNA influence	Resulting ΔT_m, ° C
M	(1+D1) - (4+D1)	+ 6.5
	(2+D1) - (5+D1)	+ 3.5
	(3+D1) - (6+D1)	+ 4.0
MS	(1+D2) - (4+D2)	+ 3.0
	(2+D2) - (5+D2)	+ 1.0
	(3+D2) - (6+D2)	+ 0.5
M	(1+R1) - (4+R1)	+ 13.0
	(2+R1) - (5+R1)	+ 9.0
	(3+R1) - (6+R1)	+ 10.0
MS	(1+R2) - (4+R2)	+ 8.0
	(2+R2) - (5+R2)	+ 6.0
	(3+R2) - (6+R2)	+ 6.0

[a] For probe sequences, see Table S1. *BRAF* targets: V600E D1: 5'-TCT AGC TAC AGA GAA ATC TCG-3' (wild type), V600E D2: 5'-TCT AGC TAC AGT GAA ATC TCG-3' (mutant), V600E R1: 5'-rUrCrU rArGrC rUrArC rA rGrA rGrArA rArUrC rUrCrG -3' (wild type), V600E R2: 5'-rUrCrU rArGrC rUrArC rArGrU rGrArA rArUrC rUrCrG -3' (mutant).

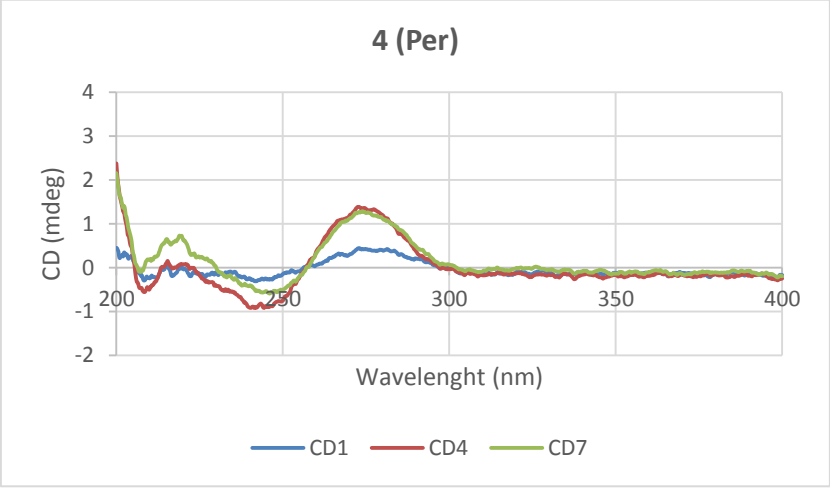
3.3. Circular dichroism measurements

CD measurements were performed using JASCO J-815 CD Spectrometer equipped with CDF4565/15 temperature controller. The solutions had a concentration of 2 μM in 1xPBS buffer. Representative spectra are shown in Figure S2.

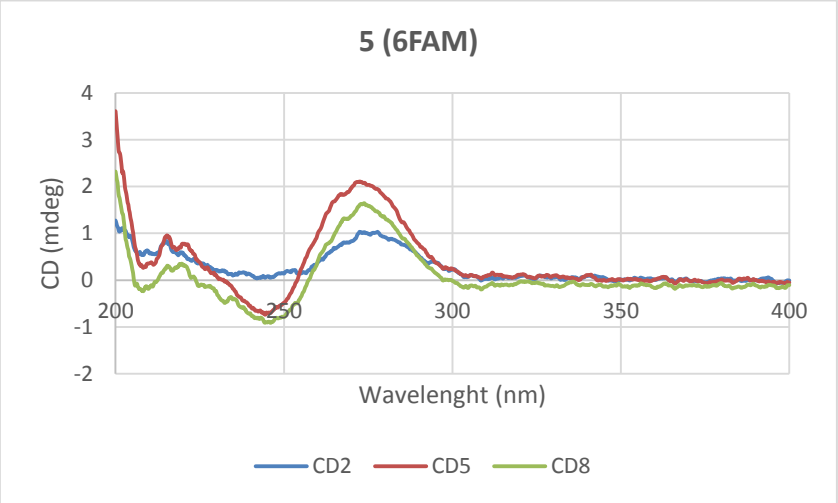
A



B



C



D

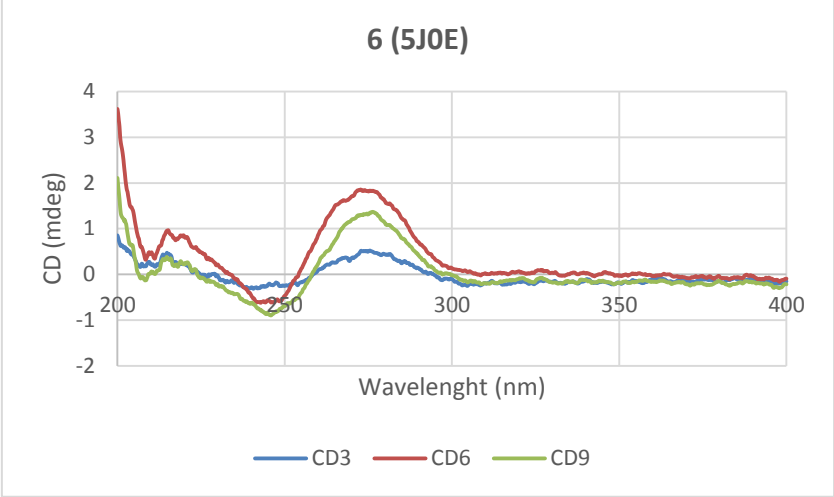


Figure S8. Circular dichroism plots of **C2** (A) and its derivatives (B-D). Blue, orange and grey lines represent the plots of single stranded probes, double stranded match and mismatch.

4. Fluorometry and fluorescence microscopy studies

4.1. Samples for fluorometry were prepared as described above (chapter **3.1**). Fluorescence measurements were made using LS 55 Luminescence Spectrometer, Perkin Elmer, equipped with a Peltier temperature controller.

The concentrations and wavelength ranges were adjusted for each dye according to its spectral parameters.[2] Thus, perylene samples had a concentration of 0.25 μM and their spectra were recorded in a wavelength range 435-645 nm. 6FAM samples had a concentration of 0.25 μM and their spectra were recorded in a wavelength range 505-645 nm. Lastly, 5JOE samples had a concentration of 0.5 μM and their spectra were recorded in a wavelength range 500-645 nm.

All the samples were prepared in 1xPBS buffer, pH 7.2.

4.2. Fluorometry results

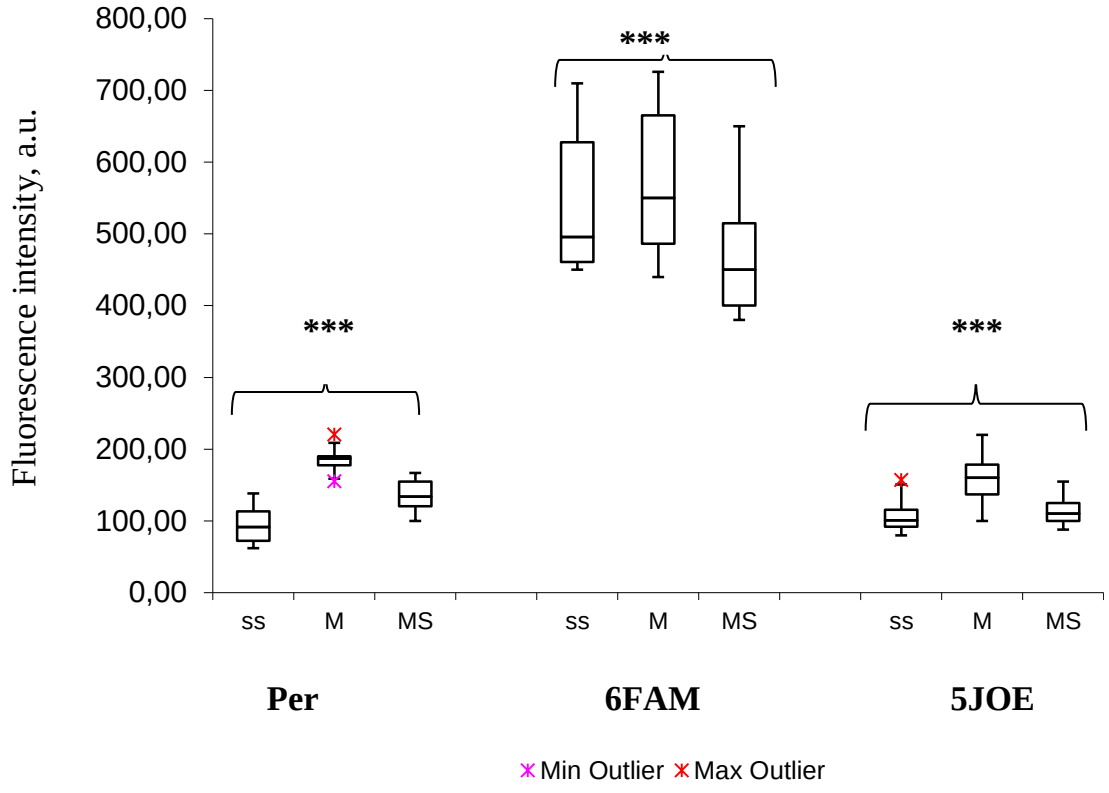


Figure S9. Box-and-whisker plot with outliers for the results of fluorometry assays with DNA targets. The arms on each boxplot are values Q1 – 1.5 × IQR and Q3 + 1.5 × IQR. Data points for each subject are means for three independent measurements. ss = single strand. M = match DNA target. MS = mismatched DNA target. Ex/em, nm: 425/455 (perylene), 505/515 (6FAM), 545/555 (5JOE); probe/target concentrations in 1xPBS, pH 7.2: 0.25 μM (perylene, 6FAM) and 0.5 μM (5JOE). *** Statistically significant difference in values at 95% significance level.

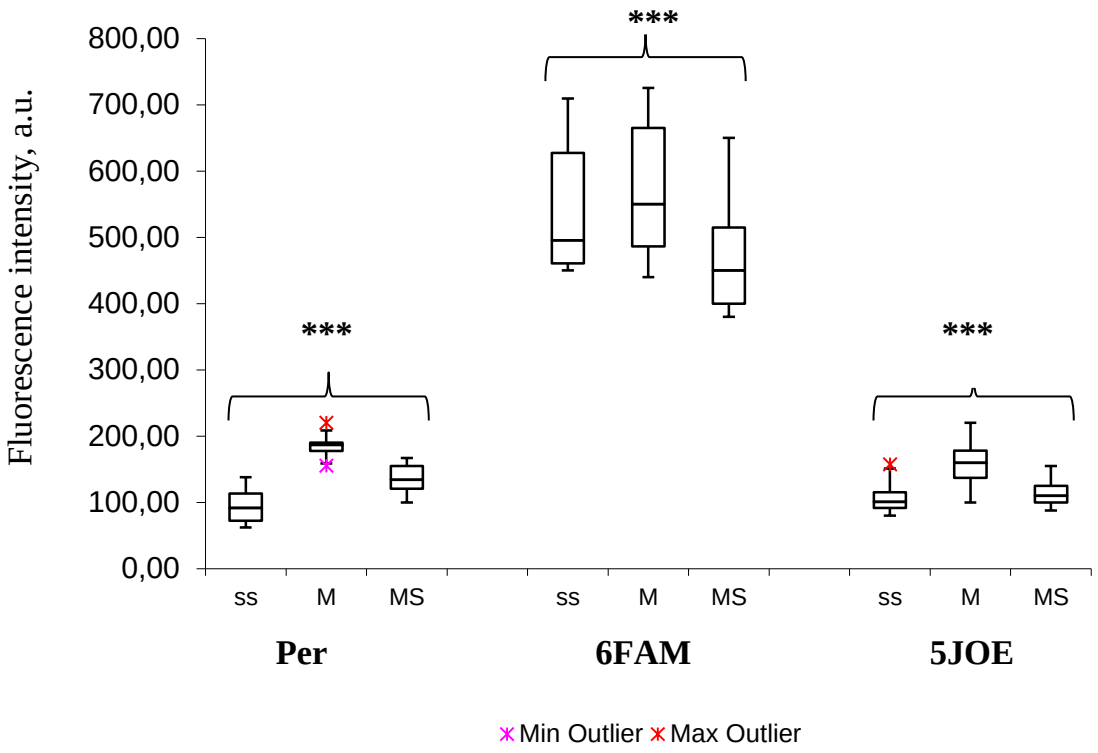


Figure S10. Box-and-whisker plot with outliers for the results of fluorometry assays with RNA targets. The arms on each boxplot are values Q1 – 1.5 × IQR and Q3 + 1.5 × IQR. Data points for each subject are means for three independent measurements. ss = single strand. M = match DNA target. MS = mismatched DNA target. Ex/em, nm: 425/455 (perylene), 505/515 (6FAM), 545/555 (5JOE); probe/target concentrations in 1xPBS, pH 7.2: 0.25 µM (perylene, 6FAM) and 0.5 µM (5JOE). *** Statistically significant difference in values at 95% significance level.

Table S8. Difference in fluorescence intensities at maxima between perylene-modified probe duplexes. M=matching; MS=mismatching; ΔF=difference in fluorescence[a]

Probe precursor	Probe no./compared complexes: M vs. MS hybridisation	ΔF
C1, 5'-CGAGA**X**TTCTCTGTAG-CX**X**AGA-3'	**1**/D1-D2	-36.90
	1/R1-R2	2.48

C2, 5'-CGAGA**X**TTC^{L}T^{L}C^{L}TGTAG C**X**AGA-3'	**4**/D1-D2	-23.06
	4/R1-R2	-31.32
C3, 5'-CGAGA**X**TTC^{L}A^{L}C^{L}TGTAG C**X**AGA-3'	**7**/D2-D1	-4.69
	7/R2-R1	-4.25

[a] For probe sequences, see Table S1. X is shown in Chart 1.
BRAF targets: V600E D1: 5'-TCT AGC TAC AGA GAA ATC TCG-3' (wild type), V600E D2: 5'-TCT AGC TAC AGT GAA ATC TCG-3' (mutant), V600E R1: 5'-rUrCrU rArGrC rUrArC rA rGrA rGrArA rArUrC rUrCrG -3' (wild type), V600E R2: 5'-rUrCrU rArGrC rUrArC rArGrU rGrArA rArUrC rUrCrG -3' (mutant). Mutation position is underlined.
Ex/em, nm: 425/455 (perylene), 505/515 (6FAM), 545/555 (5JOE); probe/target concentrations in 1xPBS, pH 7.2: 0.25 µM (perylene, 6FAM) and 0.5 µM (5JOE).

Table S9. Difference in fluorescence intensities at maxima between match and mismatch 6FAM-modified probe duplexes. M=matching; MS=mismatching; ΔF=difference in fluorescence

Probe precursor	Probe no./compared complexes: M vs. MS hybridisation	ΔF
C1, 5'-CGAGA**X**TTCTCTGTAG-C**X**AGA-3'	**2**/D1-D2	0.87
	2/R1-R2	-2.90
C2, 5'-CGAGA**X**TTC^{L}T^{L}C^{L}TGTAG C**X**AGA-3'	**5**/D1-D2	-54.49
	5/R1-R2	-17.11
C3, 5'-CGAGA**X**TTC^{L}A^{L}C^{L}TGTAG C**X**AGA-3'	**8**/D2-D1	-20.62
	8/R2-R1	19.01

Table S10. Difference in fluorescence intensities at maxima between match and mismatch 5JOE-modified probe duplexes. M=matching; MS=mismatching; ΔF=difference in fluorescence

Probe precursor	Probe no./compared complexes: M vs. MS hybridisation	ΔF
C1, 5'-CGAGA**X**TTCTCTGTAG-C**X**AGA-3'	**3**/D1-D2	-6.63
	3/R1-R2	12.20
C2, 5'-CGAGA**X**TTC^{L}T^{L}C^{L}TGTAG C**X**AGA-3'	**6**/D1-D2	7.68
	6/R1-R2	2.94
C3, 5'-CGAGA**X**TTC^{L}A^{L}C^{L}TGTAG C**X**AGA-3'	**9**/D2-D1	-0.38
	9/R2-R1	-4.66

4.3. Fluorescence microscopy[4]

Analyses were made using DNA derived from the following cell-lines: human male control HMC (contains no mutations); LS411N contains *BRAF* V600E mutation; GP2d and RCM1 contain *KRAS* G12D and G12V mutations, respectively. The fluorescence microscopy was run with a Zeiss 780 microscope supplied with GaAsP (Gallium Arsenide Phosphide) detector in a non-confocal mode. The excitation/emission wavelength was 500/540-580 nm.

Total genomic DNA from cell lines was isolated and cut in fragments of ca. 5000 bp length using endonuclease enzyme EcoRI. After the deactivation of the enzyme the DNA was purified using spin column following manufacturer's protocol (GE Life Sciences).

[4] Miotke, Astakhova, et al. PLOS One 2015

BRAF- and *KRAS*-specific 120mer capture probes (obtained from IDT) were attached by biotin-streptavidin interaction to microscopy slide. The genomic DNA was flushed onto the slide in a solution of 1X PBS buffer. The slides were incubated at 95°C for 10 min and then at 70°C for 2 hours. After that the slide was washed twice with 1xPBS buffer at 70°C. The result was single-stranded mutated cancer DNA hybridized to the 120mer capture probe.[3]

The last step was adding of the mutation specific fluorescent probe to the slide. Slides were incubated at 95°C for 10 min and then kept at 60°C for 40 min. To avoid the hybridization with mismatch strands, the slide was washed four times with 1xPBS at 45°C which is the T_m for duplex with single nucleotide mismatching DNA.

BRAF V600E enrichment probe:
5’-CAGAATTATAGAAATTAGATCTCTTACCTAAACTCTTCATAATGCTTGCTCTGATAG GAAAATGAGATCTACTGTTTTCCTTTACTTACTACACCTCAGATATATTTCTTCATGAAG ACC-biotin-3’

KRAS G12/13D enrichment probe: 5'- TCA GAA TCA TTT TGT GGA CGA ATA TGA TCC AAC AAT AGA GGT AAA TCT TGT TTT AAT ATG CAT ATT ACT GGT GCA GGA CCA TTC TTT GAT ACA GAT AAA GGT TTC TCT GAC CAT TTT CAT –biotin-3'

A **B**

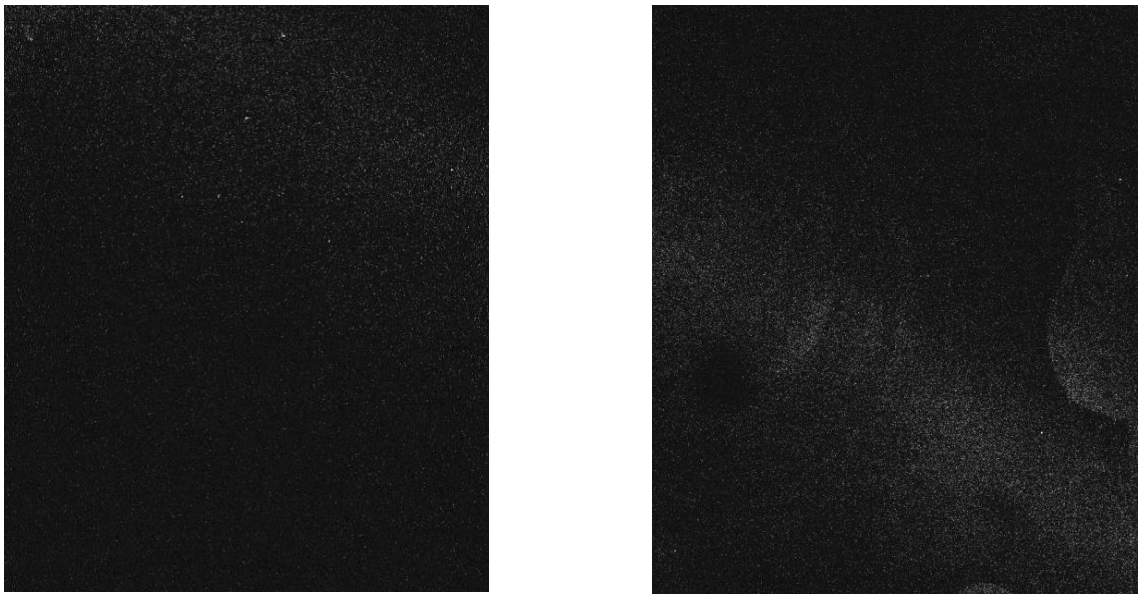


Figure S11. Negative controls: Cross-test - imaging of RCM1 DNA (contains KRAS G12V) and GP2D (contains KRAS G12D) using KRAS G12D and KRAS G13D specific probes **12** and **15**, respectively. Probe **12**: GT**M**^{3} GGAGCT^{L}G^{L}A^{L}TGGCG**M**^{3}A GGC; Probe **15**: GT**M**^{3} GGAGCT^{L}G^{L}A^{L}TGGCG**M**^{3}A GGC. Chemical structure of **M**^{3} is shown in Fig. 1.

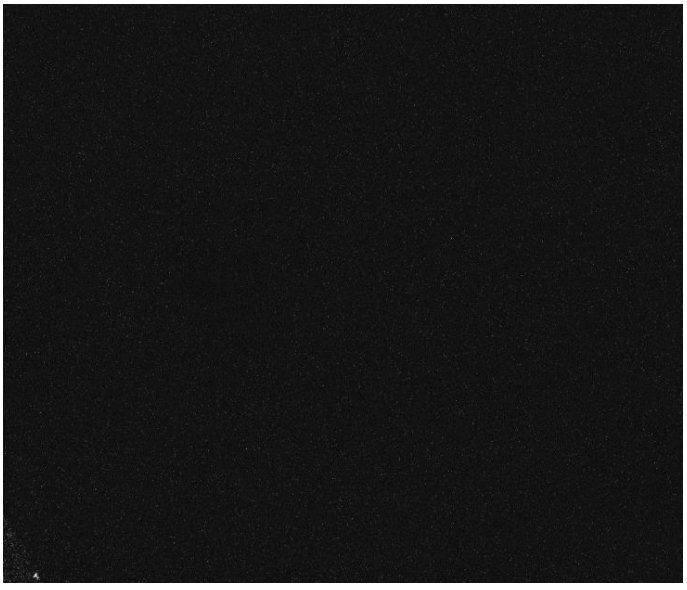


Figure S12. Negative control: Imaging of the fluorescence probe **12** without target DNA. Probe **12**: GT**M**^{3} GGAGCT^{L}G^{L}A^{L}TGGCG**M**^{3}A GGC; Chemical structure of **M**^{3} is shown in Fig. 1.

5. 5-MeC detection

T_m measurements were performed as described in chapter **3.2** (above).

Fluorometry experiments were performed as described in chapter **4** (above).

For spike-in experiment, the duplex of probe **2** with 5-MeC target (5'-TCTAG^{5Me}CTA^{5Me}CAGAGAAAT^{5Me}CT^{5Me}CG-3', 10 nM) was mixed with genomic DNA from cell line HMC and afterwards analysed by fluorometry. Each measurement was carried out twice with result deviation below 3%. Average of two measurements is given for the fluorescence intensity values.

Resulting fluorescence intensities, a.u./[HMC DNA, nM] were: <5/100; 9/50; 15/25; 22/10; 51/5.

In 5-MeC analyses, the following cell lines were used:

1. LS411N (ATCC, product no. ATCC® CRL-2159™), Organism: Homo sapiens, human, Tissue: Cecum, Disease: Dukes' Type B, Colorectal Carcinoma, 66% *BRAF* V600E
2. A-375 (Sigma, product no. 88113005), Organism: Homo sapiens, human, Tissue: Skin, Disease: Malignant Melanoma, 100% *BRAF* V600E
3. Malme-3M (ATCC, product no. ATCC® HTB-64™), Organism: Homo sapiens, human, Cell Type: Fibroblast, Tissue: Malignant Melanoma; Derived From Metastatic Site: Lung, Disease: Malignant Melanoma, 90% *BRAF* V600E[5]
4. HT-144 (ATCC, ATCC® HTB-63™), Organism: Homo sapiens, human, Cell Type: Fibroblast, Tissue: Malignant Melanoma; Derived From Metastatic Site: Subcutaneous Tissue, Disease: Malignant Melanoma, 100% *BRAF* V600E

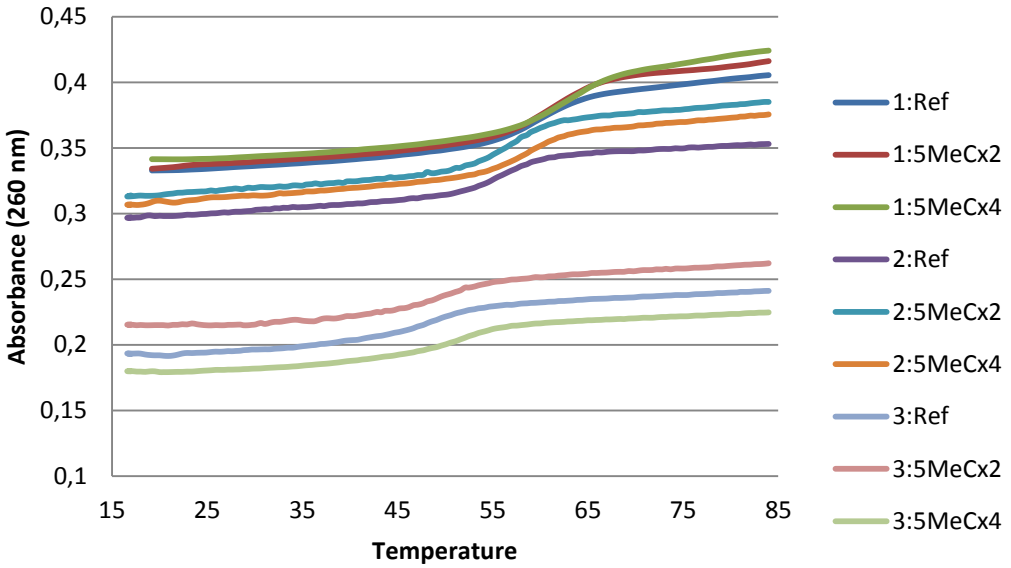
Genomic DNA preparation and solid-phase assay were carried out as described in chapter 4.3 (above).

Bisulfite DNA sequencing was carried out on Illumina HiSeq1500 instrument. DNA libraries were prepared by TruSeq DNA methylation kit (Illumina), following manufacturer's protocol.

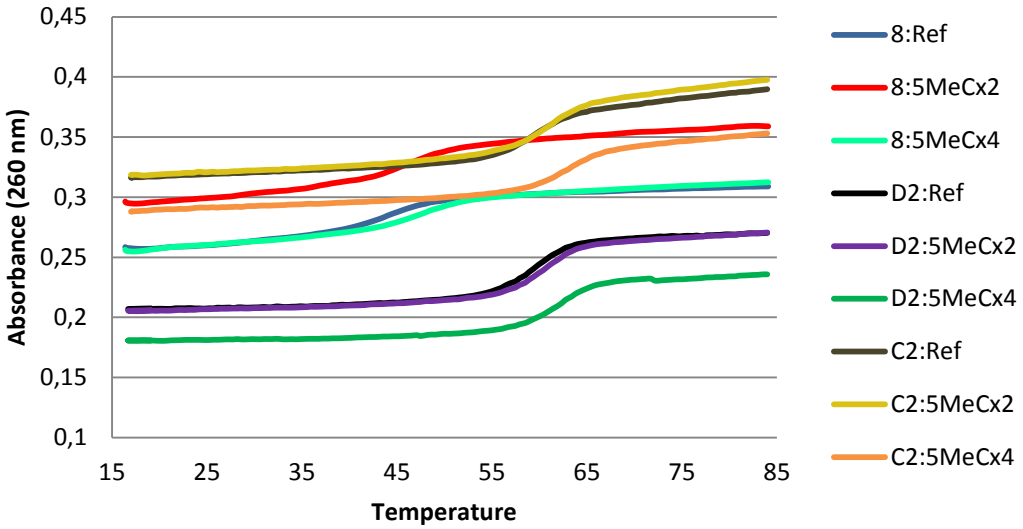
[5] Oncoimmunology 2013, 2(1): e22890.

Figure S13. Melting curves for the 5-MeC duplexes.

A



B



Probe sequences, 5'→3': **1**, CGAGA**M**^{1}TTC**T**CTGTAGC**M**^{1}AGA; **2**, CGAGA**M**^{2}TTC**T**CTGTAGC**M**^{2}AGA; **3**, CGAGA**M**^{3}TTC**T**CTGTAGC**M**^{3}AGA; **8**, CGAGA**M**^{2}TTC^{L}A^{L}C^{L}TGTAGC**M**^{2}AGA; **D2**, CGAGATTTC**T**CTGTAGCTAGA; C2, CGAGA**X**TTC^{L}**T**^{L}C^{L}TGTAGC**X**AGA. For Chemical structures of **M**^{1}-**M**^{3}, see Fig. 1 in the main paper. Monomer **X** is shown in Chart 1, SI (p. S4). LNA is marked with uppercase L. Mutation specific nucleotide within the probes is shown in red.

Target sequences, 5'→3': **Ref:** TCTAGCTACAGAGAAATCTCG; **5MeCx2**: TCTAG^{5Me}CTACAGAGAAATCT^{5Me}CG, **5MeCx4**: TCTAG^{5Me}CTA^{5Me}CAGAGAAAT^{5Me}CT^{5Me}CG. Methylated dC is underlined.

Table S11. Sequencing results for *BRAF* V600E region.

Cell line	Sequence, 5'→3', 1788-1808 (1799T>A)	% 5-MeC	Error
LS411N	TCTAGCTACAG**A**GAAATCT^{5Me}CG	16	±1%
A-375	TCTAGCTACAG**A**GAAATCT^{5Me}CG	14	±1%
Malme-3M	TCTAGCTACAG**A**GAAATCT^{5Me}CG	8	±1%
HT-144	TCTAGCTACAG**A**GAAATCT^{5Me}CG	4	±1%
HMC (negative control)	no signal	-	-

6. Statistical analyses

The statistical analyses were done using one-way ANOVA in R.[6]

[6] R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/