Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Nicked Invader probes: Multistranded and sequence-unrestricted recognition of double-stranded DNA

Shiva P. Adhikari, Saswata Karmakar and Patrick J. Hrdlicka*

Department of Chemistry, University of Idaho, Moscow, ID-83844, USA

E-mail: hrdlicka@uidaho.edu

Electronic Supplementary Information

Definition of zipper nomenclature. The following nomenclature is used to describe the relative arrangement between two monomers on opposing strands in a duplex. The number *n* describes the distance measured in number of base-pairs and has a positive value if a monomer is shifted toward the 5′-side of its strand relative to a second reference monomer on the other strand. Conversely, *n* has a negative value if a monomer is shifted toward the 3′-side of its strand relative to a second reference monomer on the other strand.

Table S1. MS data of ONs used in this study.[a]

ON	Sequence	Calculated *m/z* $(M+H)^+$	Observed *m/z* $(M+H)^+$
1[b]	5'-**U**GCA**C**AGGTA**U**A**U**ATAGGC	6731.0	6731.5
2	5'- **C**GCA**U**A'	2223.0	2222.0
3	3'-A**C**GTG**U**	2254.0	2253.0
4[b]	3'- CCATA**U**A**U**ATCCGG**C**GTA**U**	6642.0	6642.5
5	5'-**U**GCA**C**AGGTA**U**A**U**ATAG	6112.5	6113.0
6	5'-GC**C**GCA**U**A	2841.0	2841.5
7	3'-A**C**GTG**U**CC	2832.0	2833.5
8	3'-ATA**U**A**U**ATCCGG**C**GTA**U**	6063.5	6064.5
9	5'-**U**GCA**C**AGGTA**U**A**U**AT	5470.0	5470.0
10	5'-AGGC**C**GCA**U**A	3483.5	3484.0
11	5'-A**C**GTG**U**CCAT	3449.5	3449.5
12	3'-A**U**A**U**ATCCGG**C**GTA**U**	5446.0	5447.0
13	5'-**U**GCA**C**AGGTA**U**A**U**	4852.5	4853.0
14	5'-ATAGGC**C**GCA**U**A	4101.0	4101.5
15	3'-A**C**GTG**U**CCATA**U**	4283.0	4283.0
16	5'-A**U**ATCCGG**C**GTA**U**	4612.5	4613.0
17	5'-GGTA**U**A**U**ATAGGC**C**GCA**U**A	6731.0	6732.0
18	3'-A**C**GTG**U**CCATA**U**A**U**ATCCG	6265.5	6266.0
19	5'-GGTA**U**A**U**ATAGGC	4446.0	4447.5
20	3'-CCATA**U**A**U**ATCCG	4326.0	4327.0
21	5'-Cy3-T**U**A**U**ATGCTG**U**TCTC	5680.0	5678.0
22	3'-AA**U**A**U**ACGACA**A**GAG-Cy3	5789.0	5787.0
23	5'-Cy3-TG**U**GT**U**ATA**U**GC**U**GTTC	6570.0	6570.0
24	5'-T**C**AGCC**C**T	2806.0	2806.0
25	3'-ACA**C**AA**U**A	2833.0	2833.0
26	3'-TA**C**GA**C**AAGAG**U**CGGG**A**-Cy3	6683.0	6681.5

[a] MALDI-MS used to obtain data for all ONs except **2**, **3** and **23**-**26** for which LC-ESI-MS was used. Data for **1**, **4** and **17**-**20** are from reference S1. Data for **21** and **22** are from reference S2. For MALDI-MS, the *m/z* is for the $(M+H)^+$ ion, while it is for the $(M-H)^-$ ion with LC-ESI-MS.

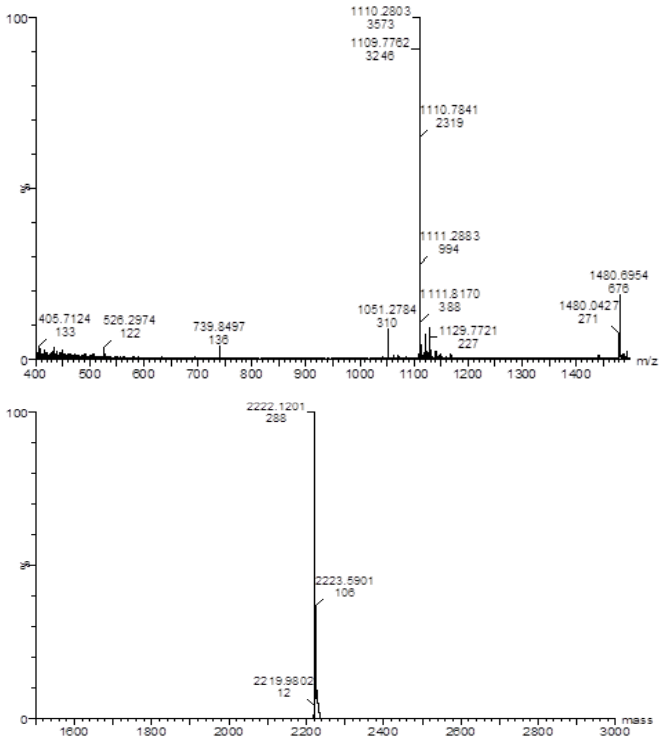


Figure S1. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON2**.

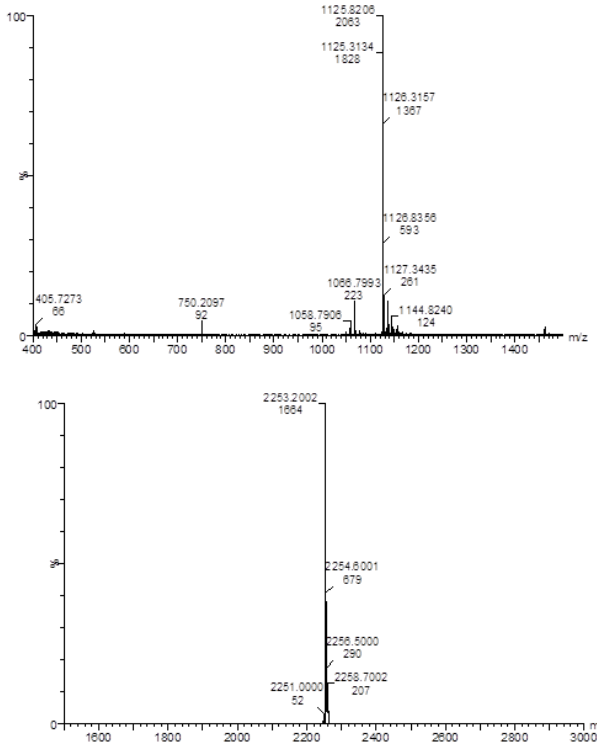


Figure S2. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON3**.

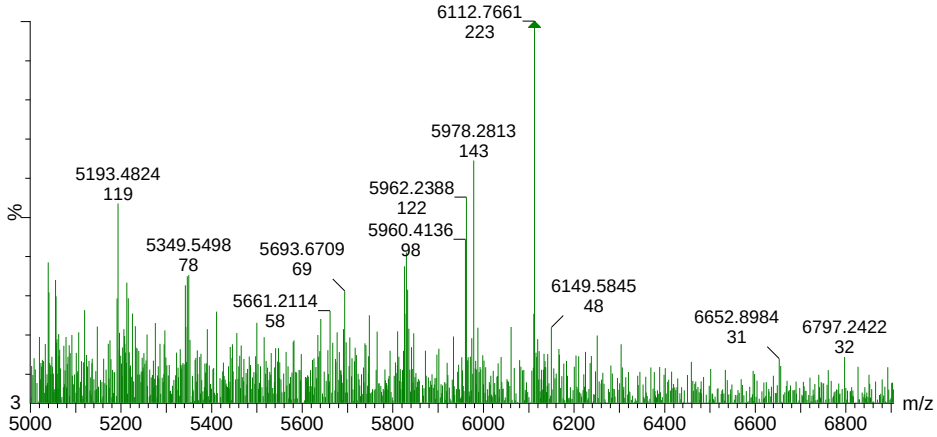


Figure S3. MALDI-MS spectrum of **ON5**.

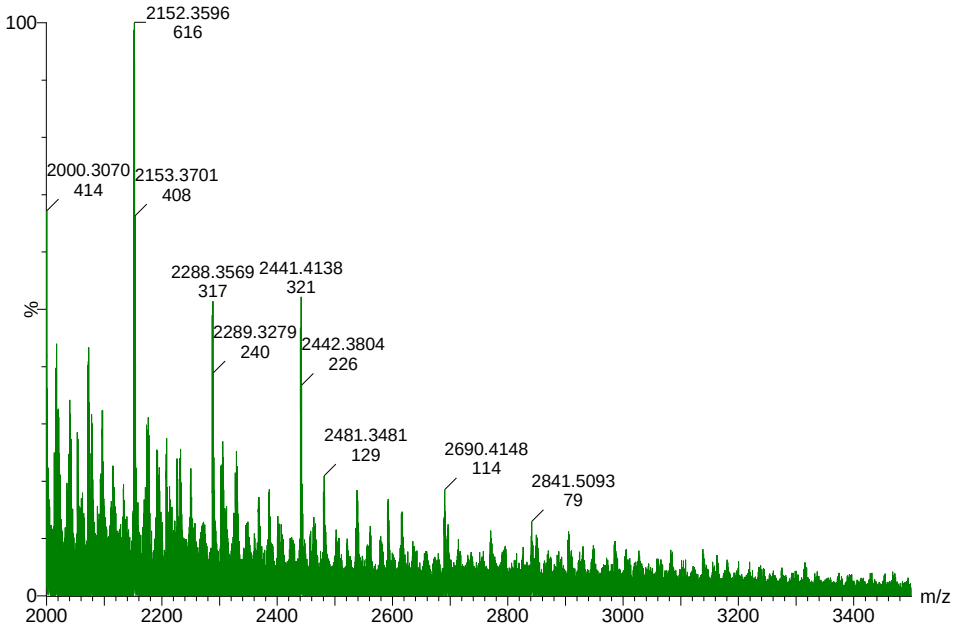


Figure S4. MALDI-MS spectrum of **ON6**.

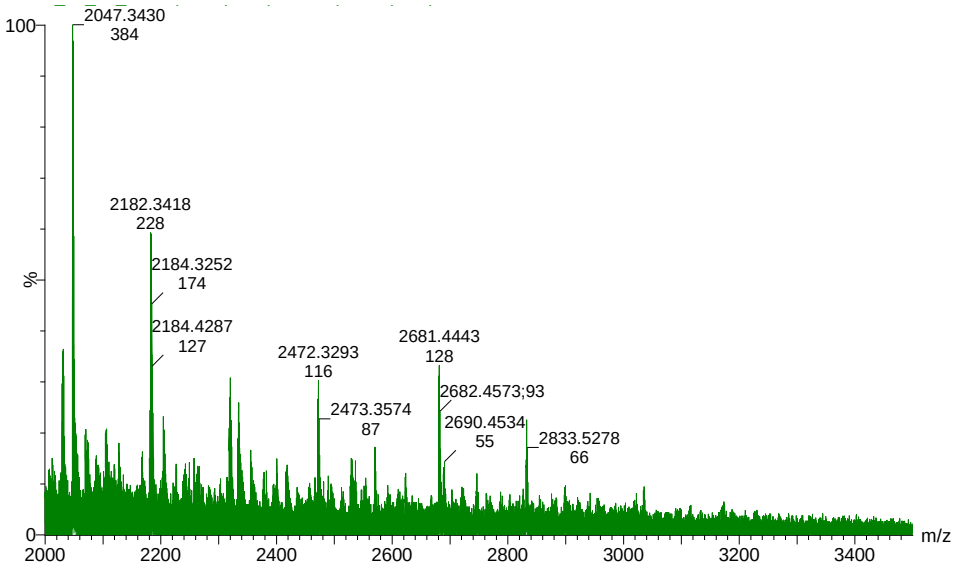


Figure S5. MALDI-MS spectrum of **ON7**.

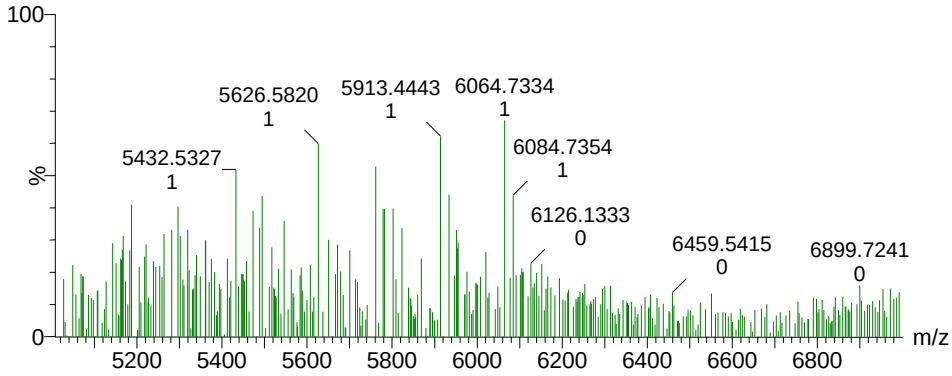


Figure S6. MALDI-MS spectrum of **ON8**.

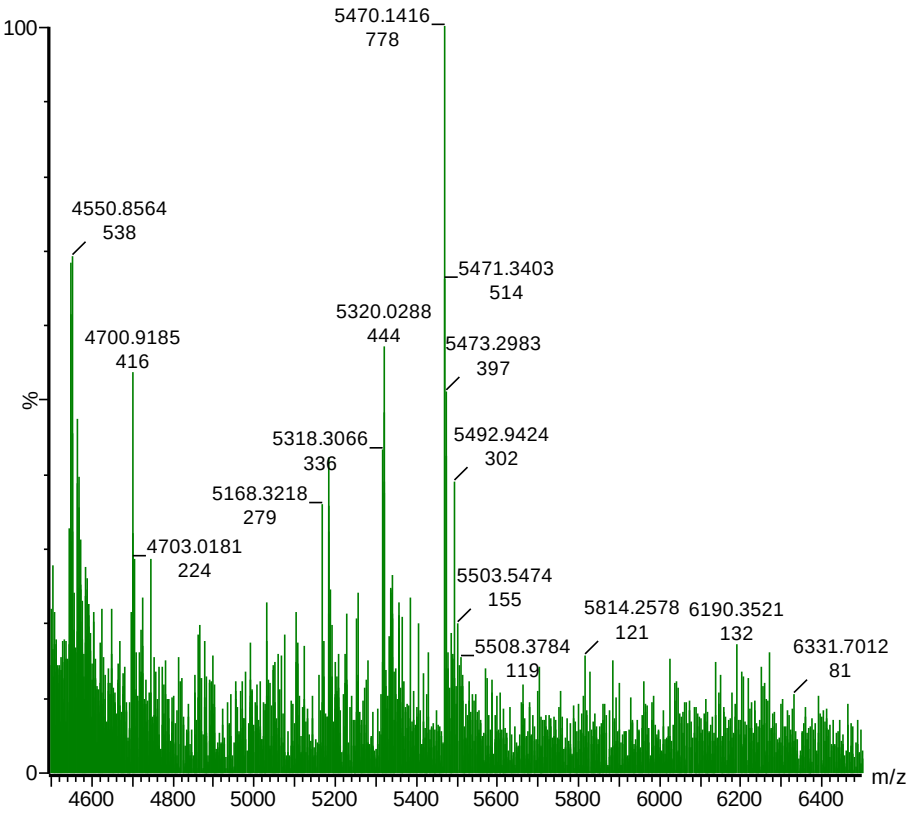


Figure S7. MALDI-MS spectrum of **ON9**.

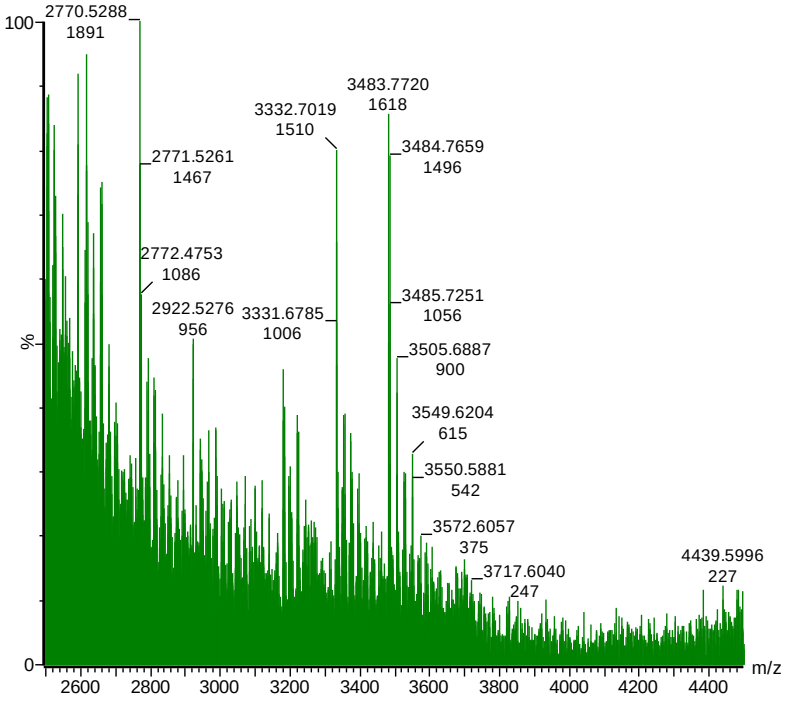


Figure S8. MALDI-MS spectrum of **ON10**.

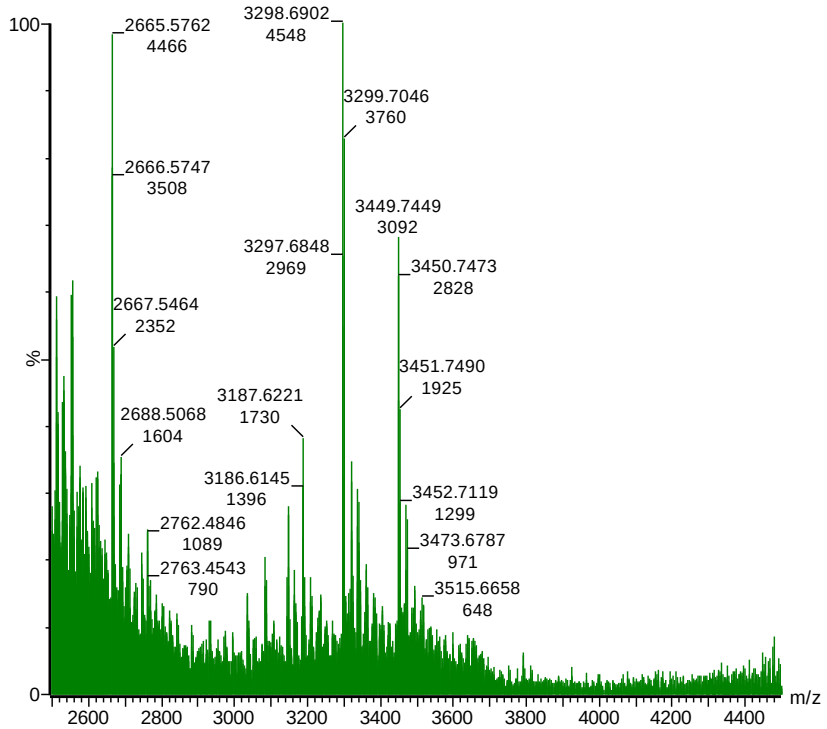


Figure S9. MALDI-MS spectrum of **ON11**.

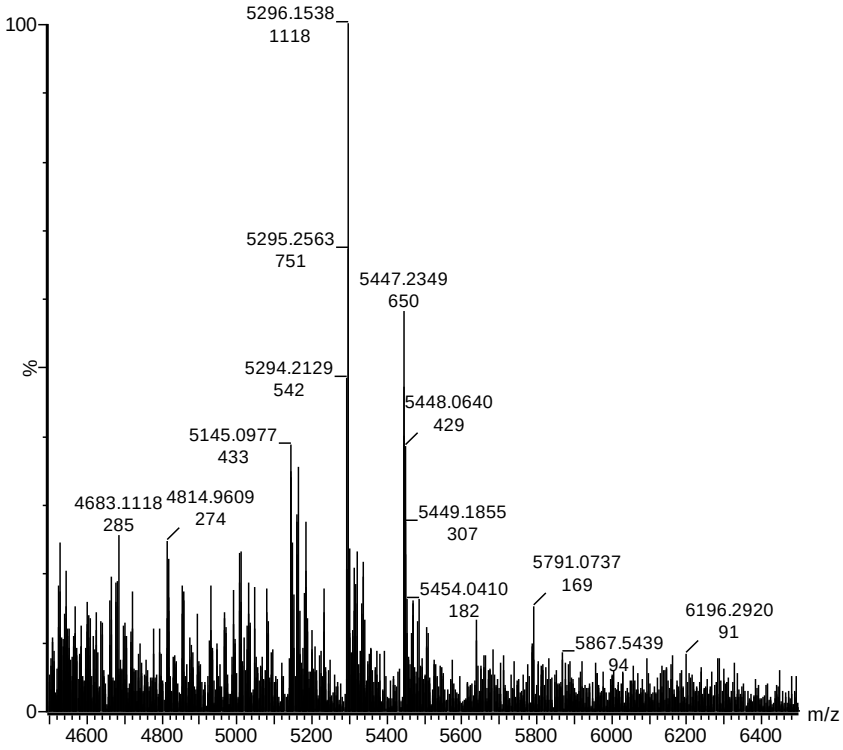


Figure S10. MALDI-MS spectrum of **ON12**.

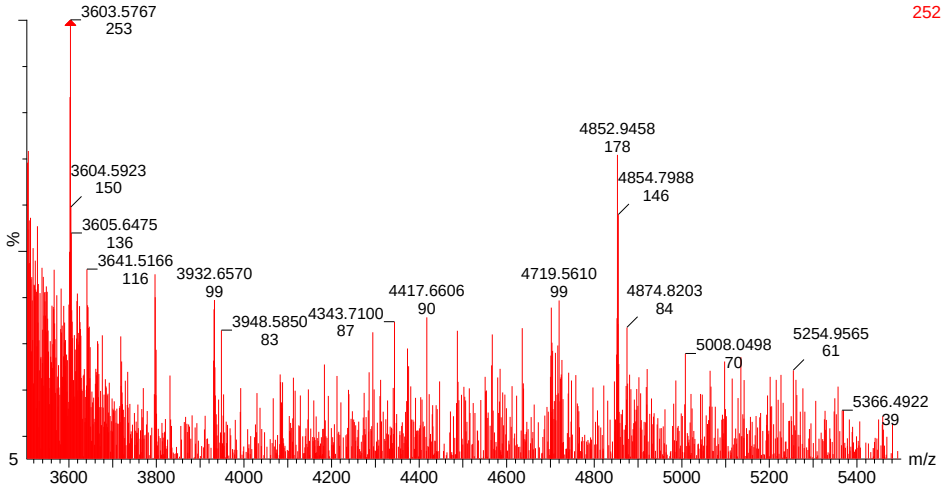


Figure S11. MALDI-MS spectrum of **ON13**.

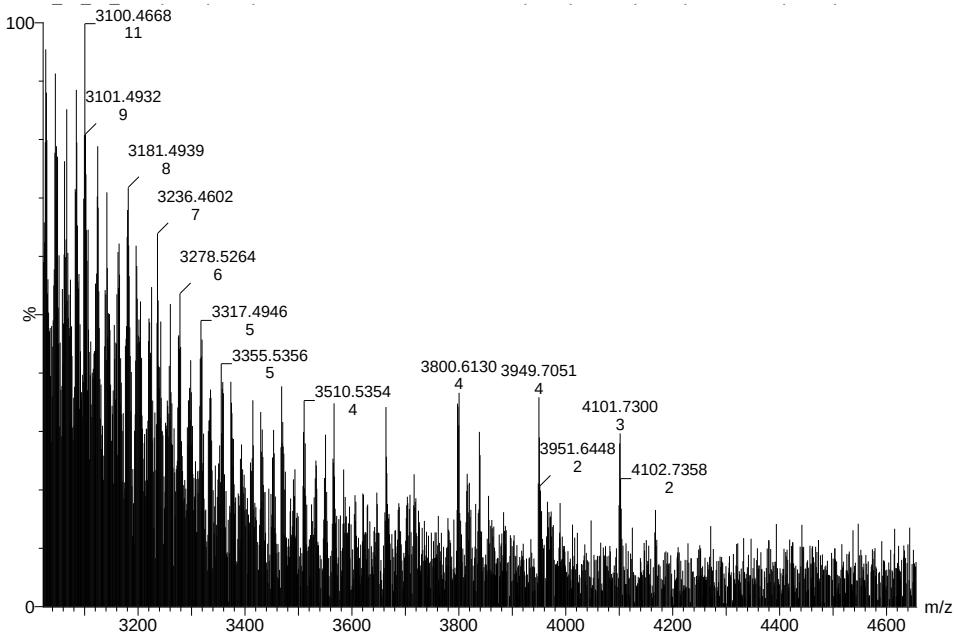


Figure S12. MALDI-MS spectrum of **ON14**.

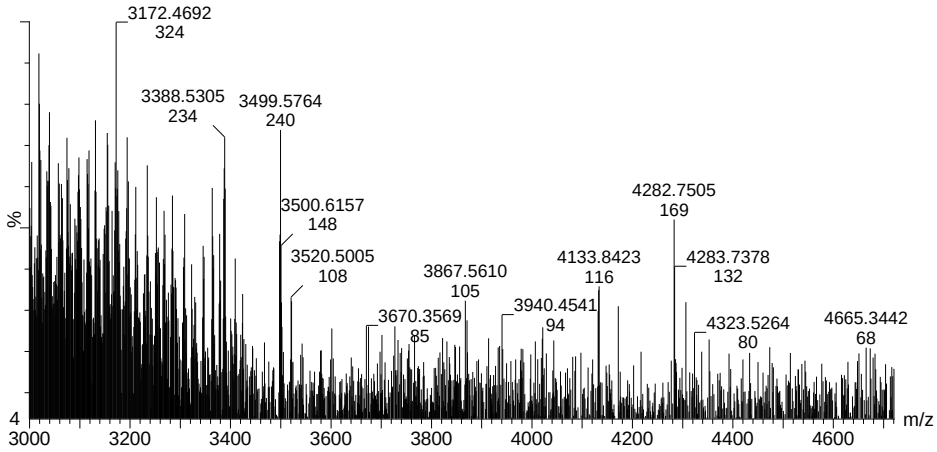


Figure S13. MALDI-MS spectrum of **ON15**.

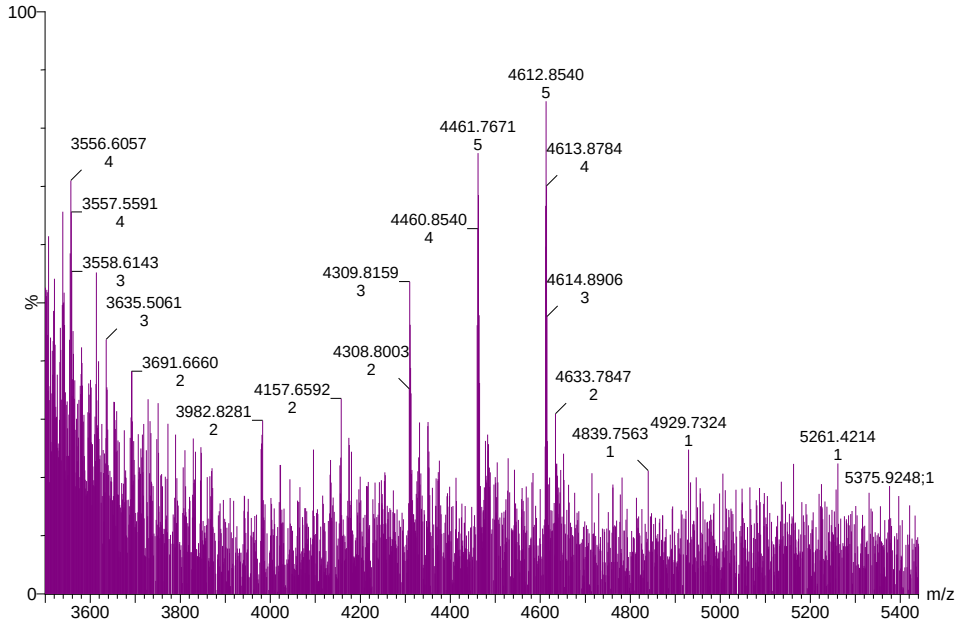


Figure S14. MALDI-MS spectrum of **ON16**.

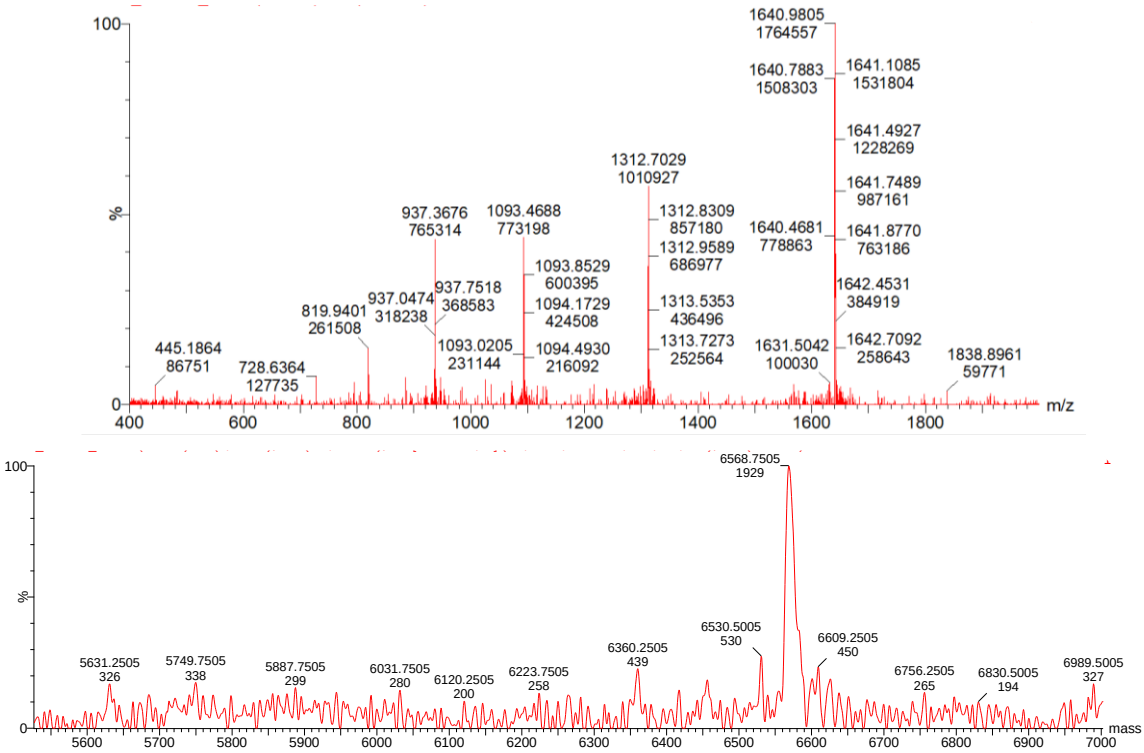


Figure S15. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON23**.

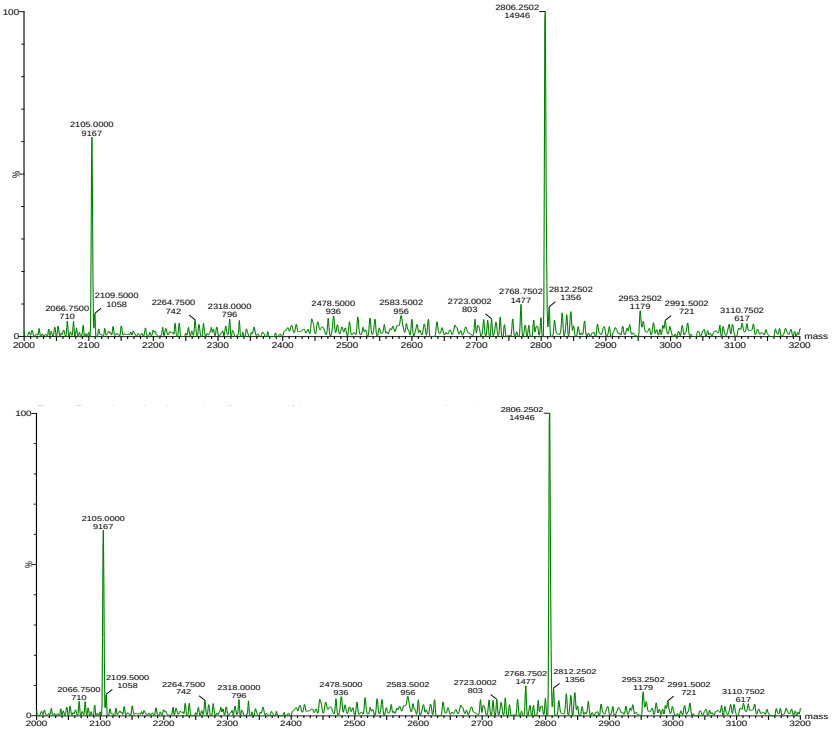


Figure S16. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON24**

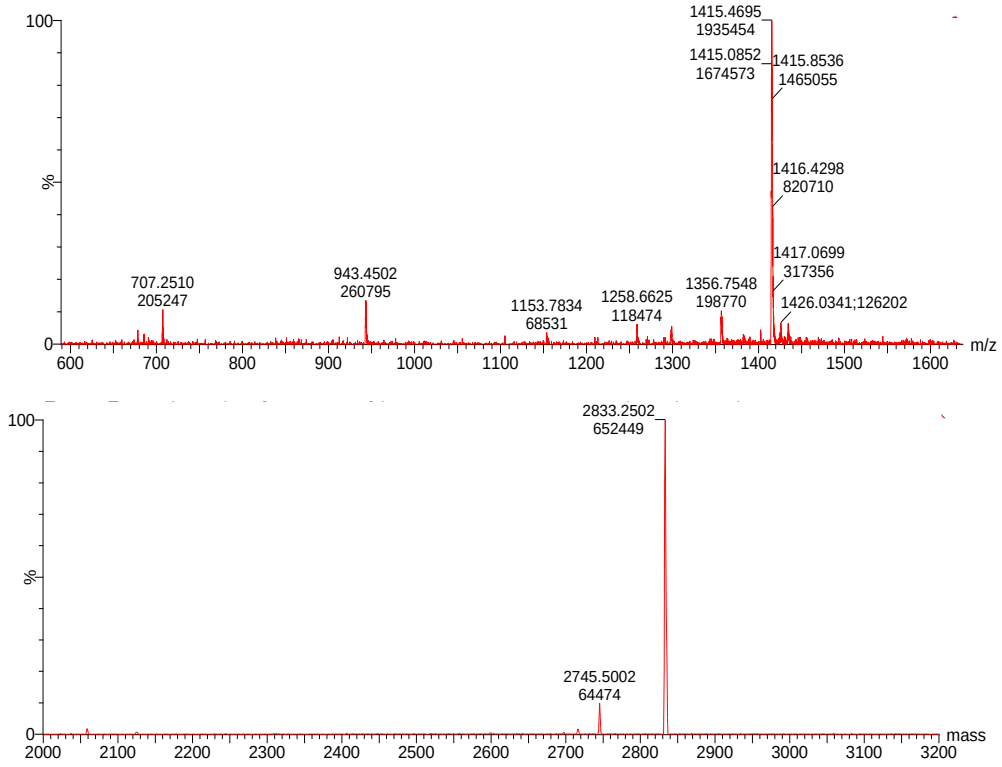


Figure S17. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON25**.

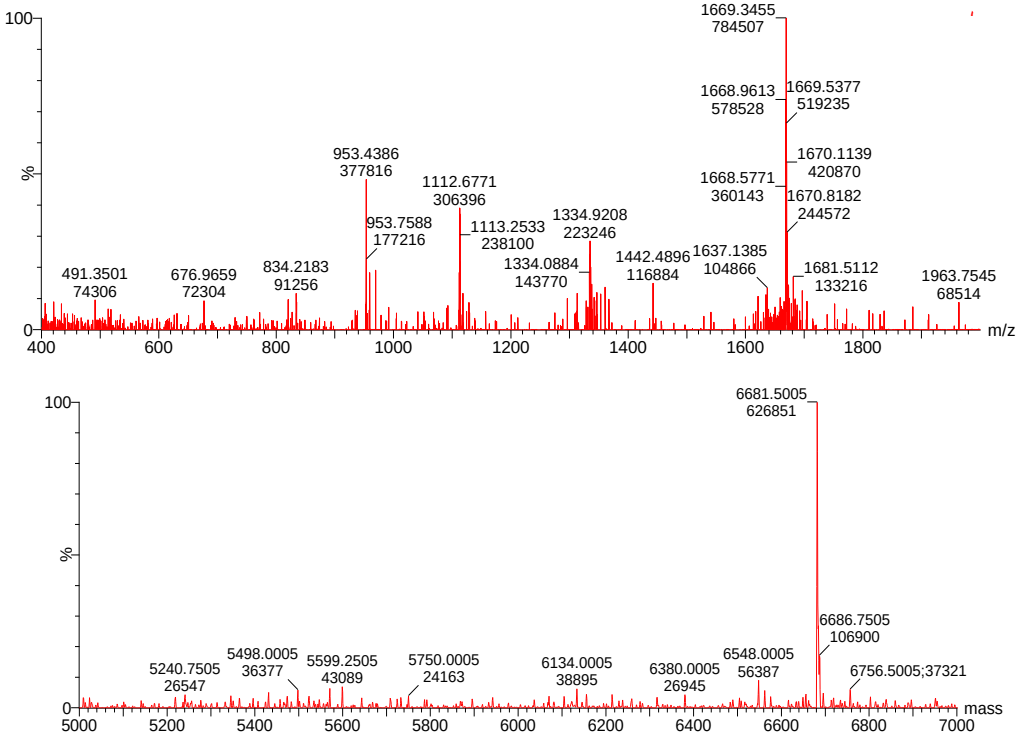
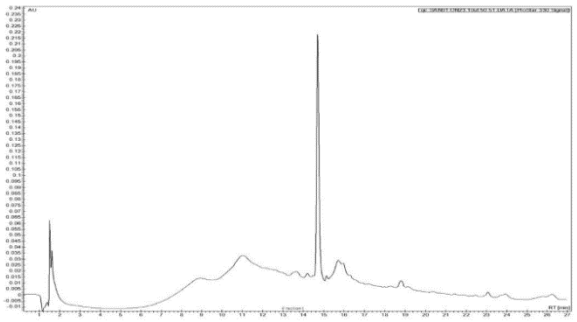
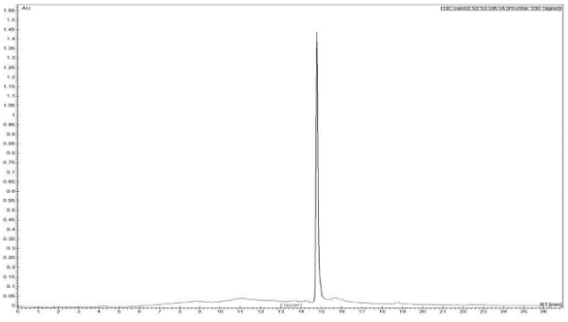


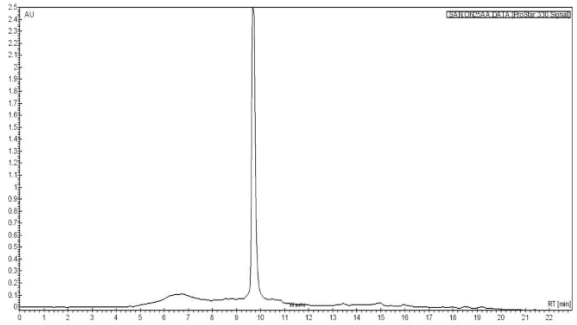
Figure S18. Unprocessed (upper panel) and deconvoluted (lower panel) ESI-MS spectrum of **ON26**.



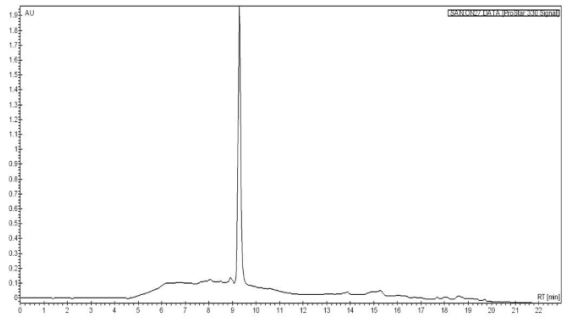
ON2



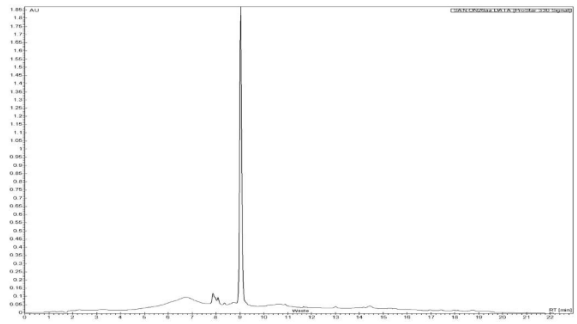
ON3



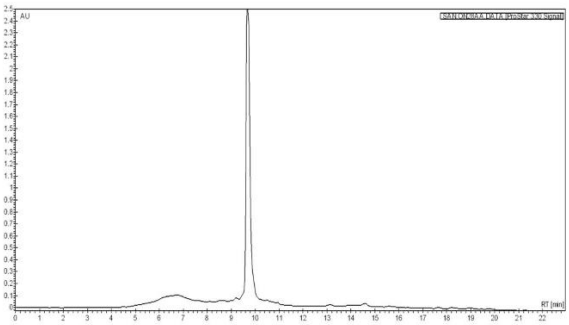
ON5



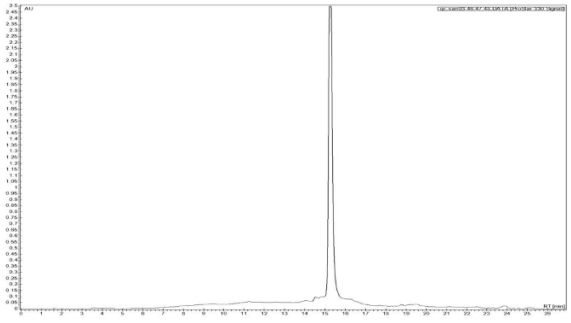
ON6



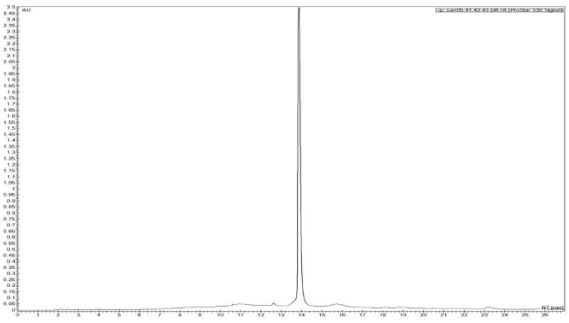
ON7



ON8

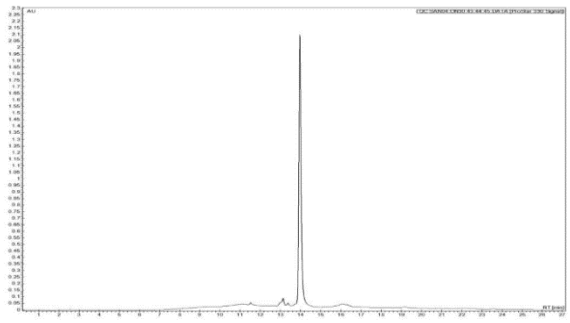


ON9

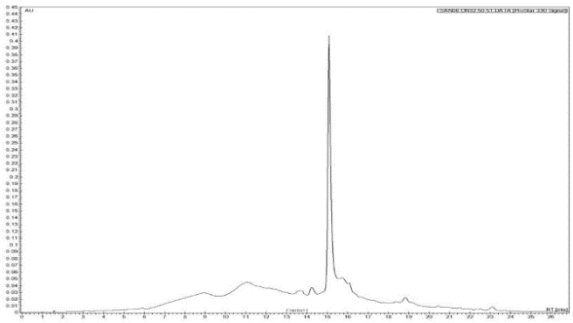


ON10

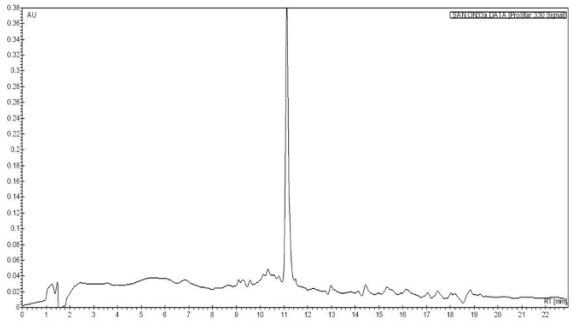
Figure S19. HPLC traces of ONs used in this study.



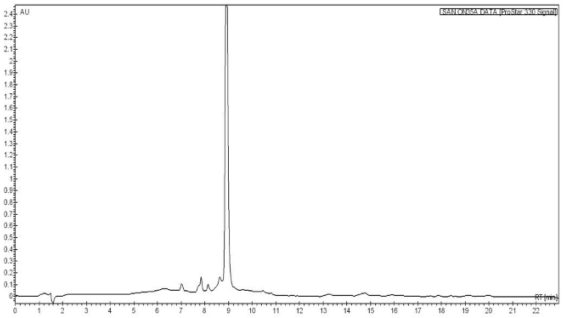
ON11



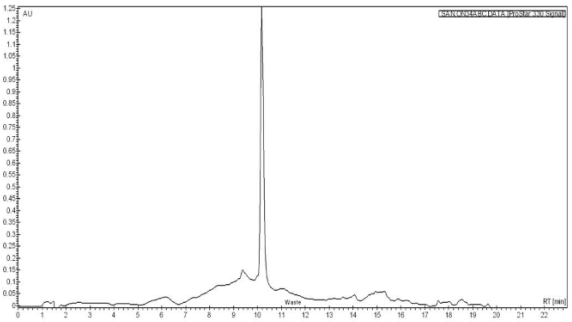
ON12



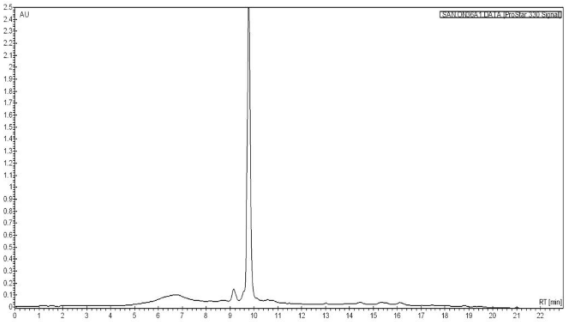
ON13



ON14



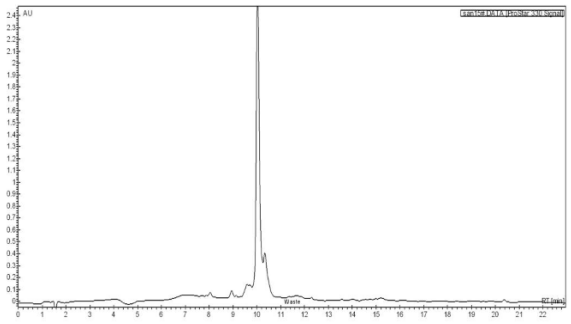
ON15



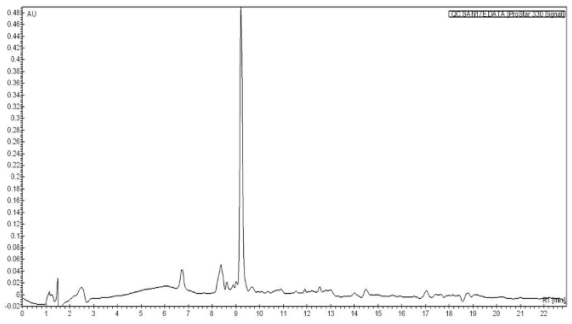
ON16

.

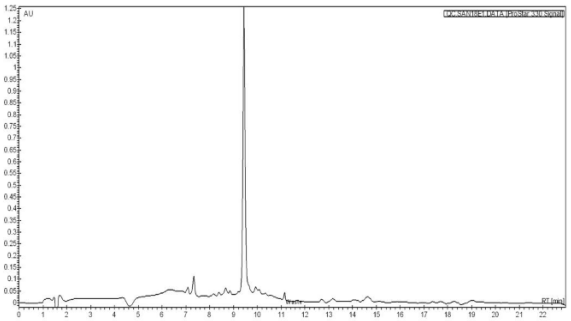
Figure S20. HPLC traces of ONs used in this study.



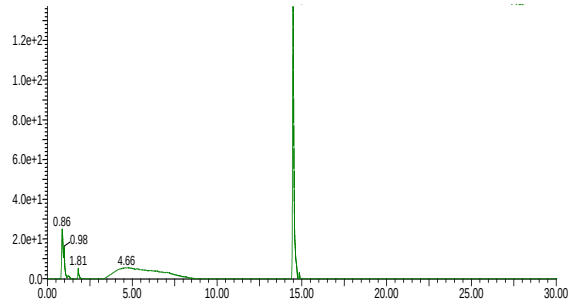
ON23



ON24



ON25



ON26

Figure S21. HPLC traces of ONs used in this study. Chromatogram for **ON26** obtained from LC-ESI-MS analysis.

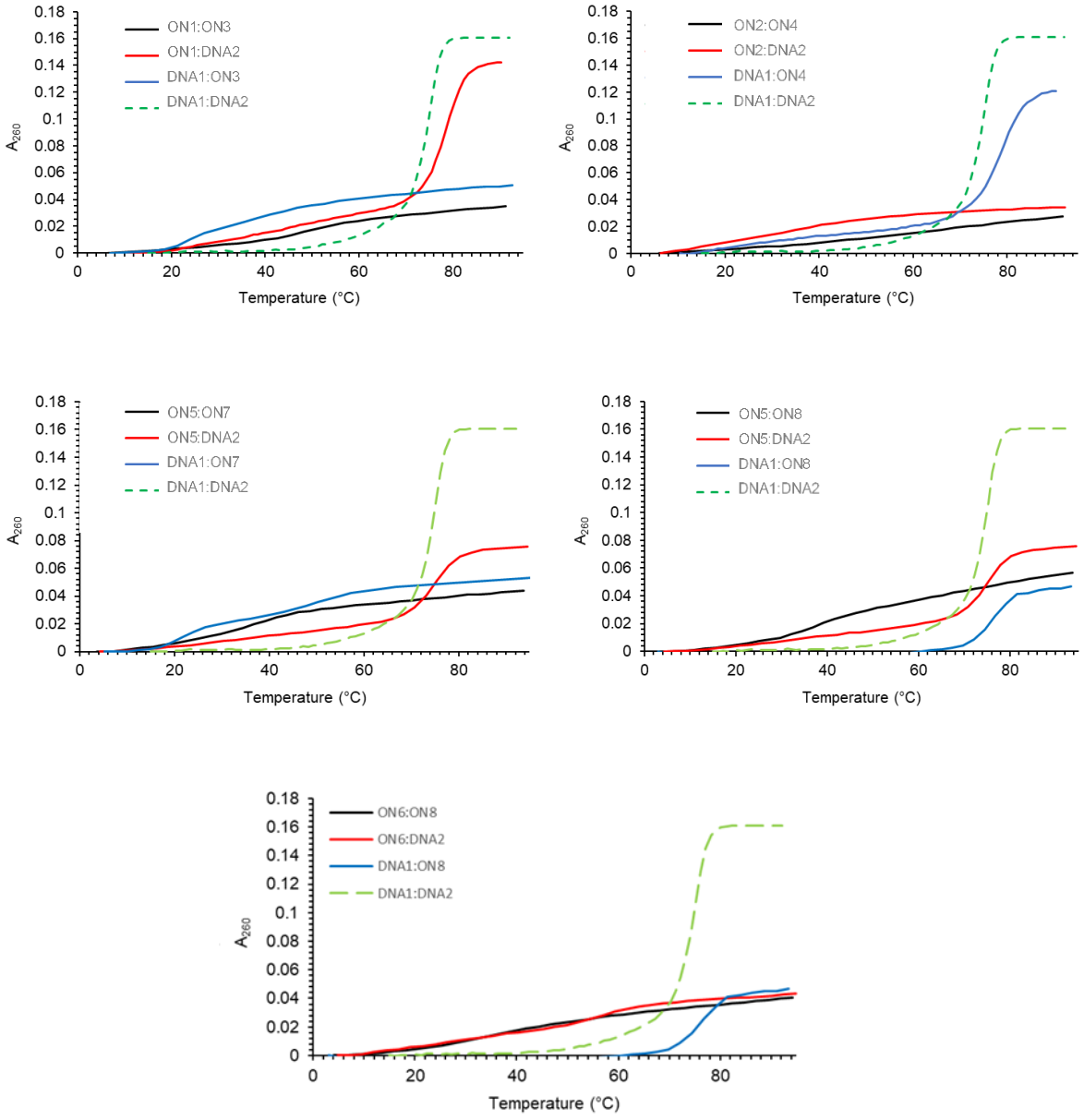


Figure S22. Representative thermal denaturation profiles of double-stranded probes, duplexes between individual probe strands and 33-mer ssDNA harboring complementary target regions, and unmodified reference duplex **DNA1**:**DNA2**. For experimental conditions, see Table 1.

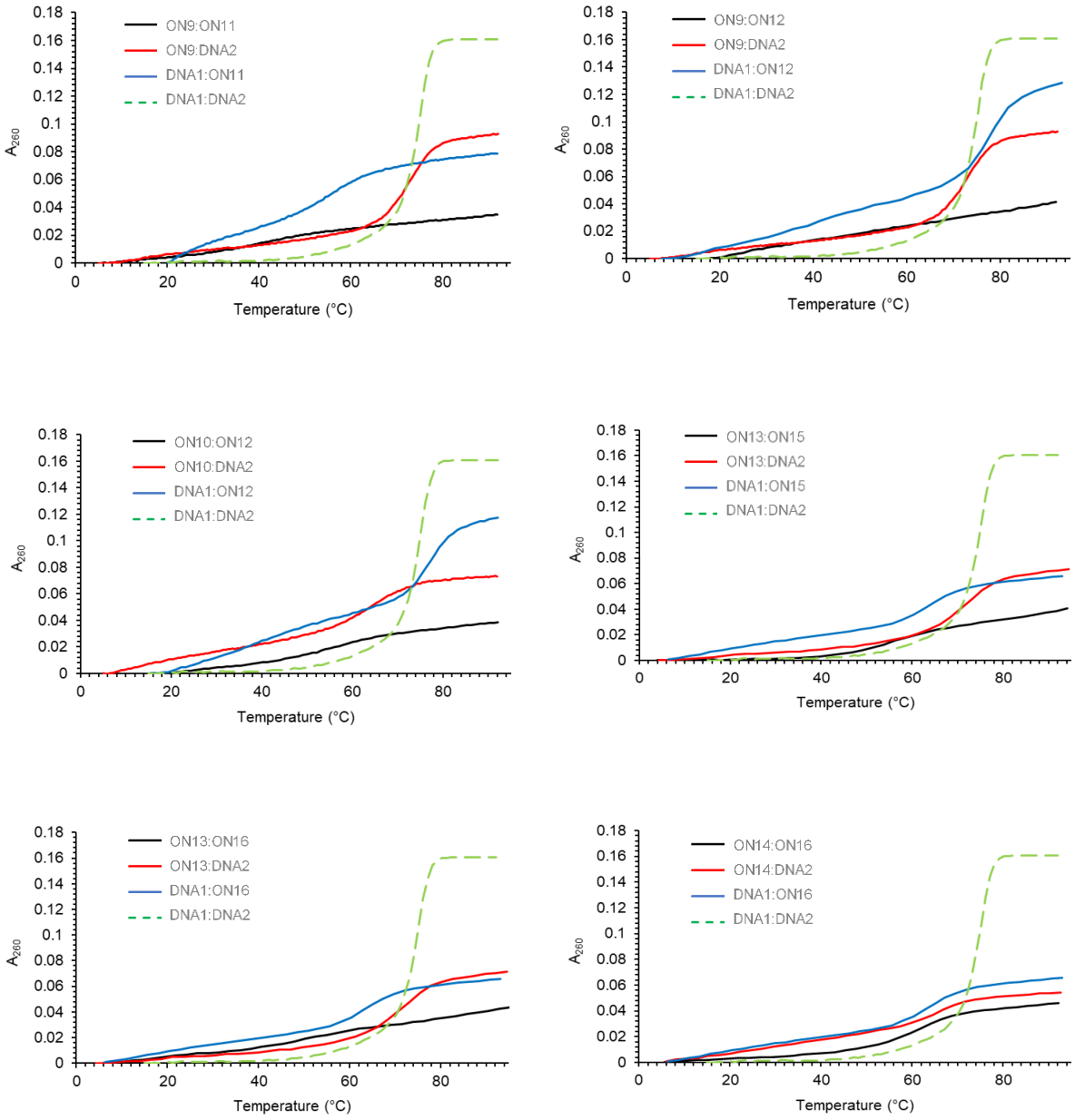


Figure S23. Representative thermal denaturation profiles of double-stranded probes, duplexes between individual probe strands and 33-mer ssDNA harboring complementary target regions, and unmodified reference duplex **DNA1**:**DNA2**. For experimental conditions, see Table 1.

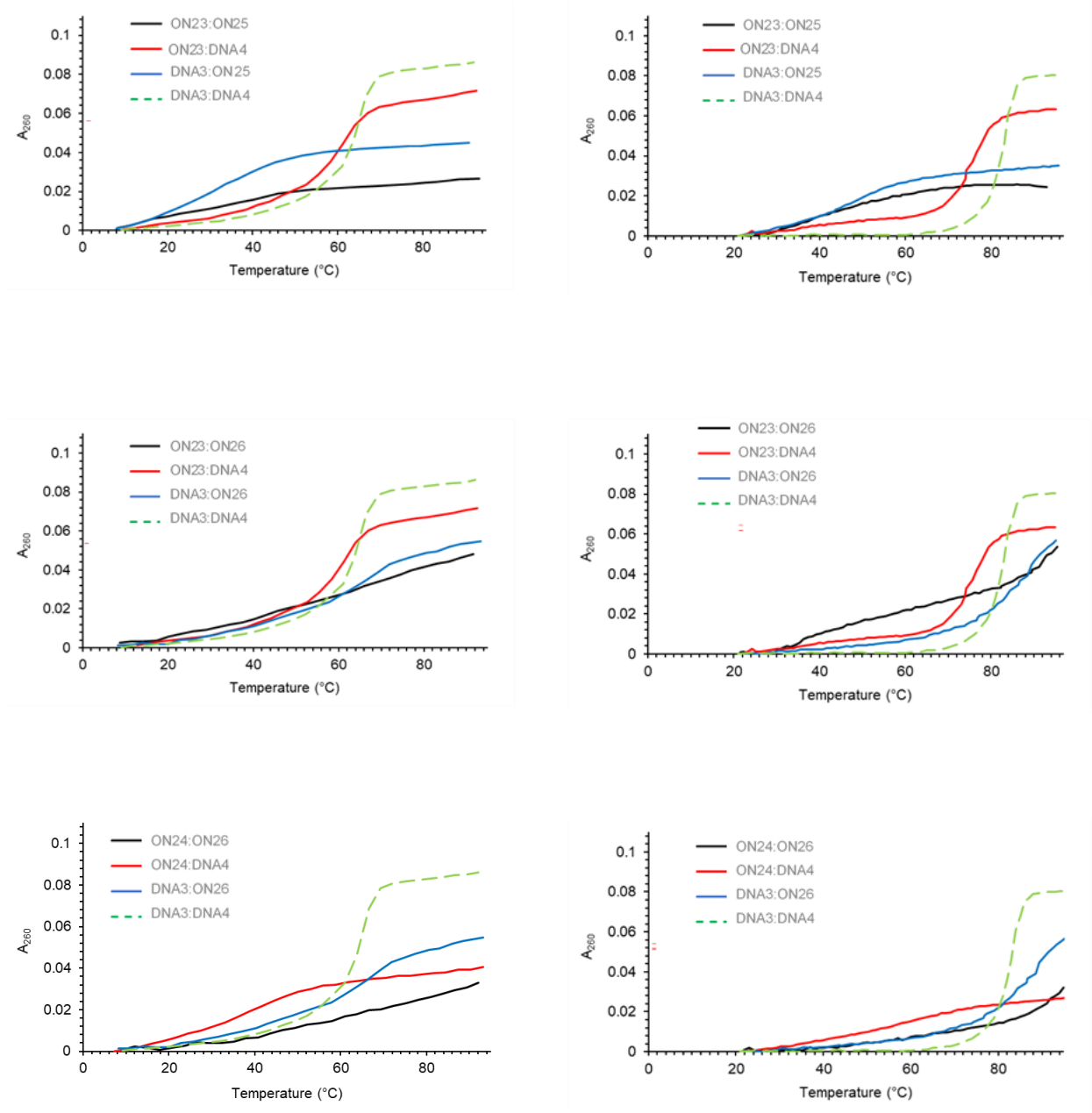


Figure S24. Representative thermal denaturation profiles of probes used in FISH experiments (**DYZ-NIP** = **ON23**/**24**:**ON25**/**26**), duplexes between individual probe strands and 33-mer ssDNA harboring complementary target regions, and reference duplex **DNA3**:**DNA4** recorded in medium (left) or high (right) salt buffer. For experimental conditions, see Table 1.

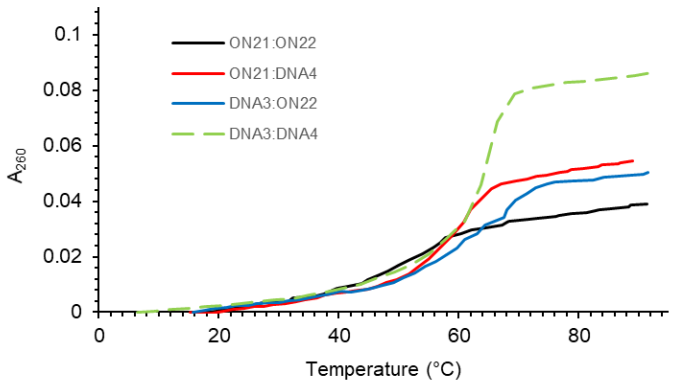


Figure S25. Representative thermal denaturation profiles of probes used in FISH experiments (**DYZ-REF**), duplexes between individual probe strands and 33-mer ssDNA harboring complementary target regions, and reference duplex **DNA3**:**DNA4** recorded in medium buffer. For experimental conditions, see Table 1.

Supplemental discussion regarding denaturation profiles and T_m values. While evaluating the stability of the different duplex segments of **NIP1**, we observed a transition for **ON1**:**ON3** at a temperature of ~45 °C. Based on our prior experience with these type of probes, this transition seemed unlikely to stem from melting of a 6-bp duplex with two energetic hotspots (the single-stranded overhangs were expected only to have a minor influence on the T_m value).[S1] Indeed, when recording the thermal denaturation profile for **ON1** in absence of **ON3**, a transition at a similar temperature was observed prompting us to suggest that the ~45 °C transition likely is due to denaturation of a secondary structure entailing only **ON1** (Fig. S26). No other transitions were observed for **ON1**:**ON3**, prompting us to list "no transition" for **ON1**:**ON3** in Table 1.

ON1:**ON4** displays a transition at 49 °C which is close to the T_m value observed for **ON1**-only (Fig. S26a). To determine if this transition is due to duplex formation between **ON1** and **ON4** or denaturation of a secondary structure entailing **ON1**-only, we recorded denaturation profiles for **ON1**:**ON4** and **ON1**-only at 10-fold higher strand concentrations. In both instances, denaturation curves were shifted to a higher, identical T_m value (~57 °C, Fig. S26b and c). A differential thermal denaturation curve (i.e., **ON1**:**ON4** – **ON1**-only) also displayed a T_m value of ~57 °C, seemingly suggesting concomitant heteroduplex (**ON1**:**ON4**) and homoduplex (**ON1**-only) formation when **ON1** and **ON4** are mixed. Further evidence supporting duplex formation between **ON1** and **ON4** was obtained from electrophoretic mobility shift assays conducted at non-denaturing conditions, as slower moving bands were observed when **ON1** and **ON4** were mixed (Fig. S27).

The denaturation profiles of **DNA1**:**ON7** and **DNA1**:**ON11** display two transitions (Fig. S28a and b). Analysis of the 33-mer **DNA1** strand using the OligoAnalyzer tool from Integrated DNA Technologies, hinted at possible intramolecular hairpin formation with an estimated T_m of ~25 °C. Indeed, denaturation profiles of **DNA1**-only suggest the presence of a transition below 30

°C (see Fig. S28a and b). Thus, it is reasonable to attribute the first transition to the denaturation of this secondary structure, and the second transition to the denaturation of the probe-target duplexes.

In light of the above and to clarify the nature of the transition observed at ~25 °C for **DNA1**:**ON3**, we constructed a differential thermal denaturation curve (i.e., **DNA1**:**ON3** - **DNA1**), which suggests that **DNA1**:**ON3** forms a labile duplex (T_m ~25.5 °C, Fig. S28c).

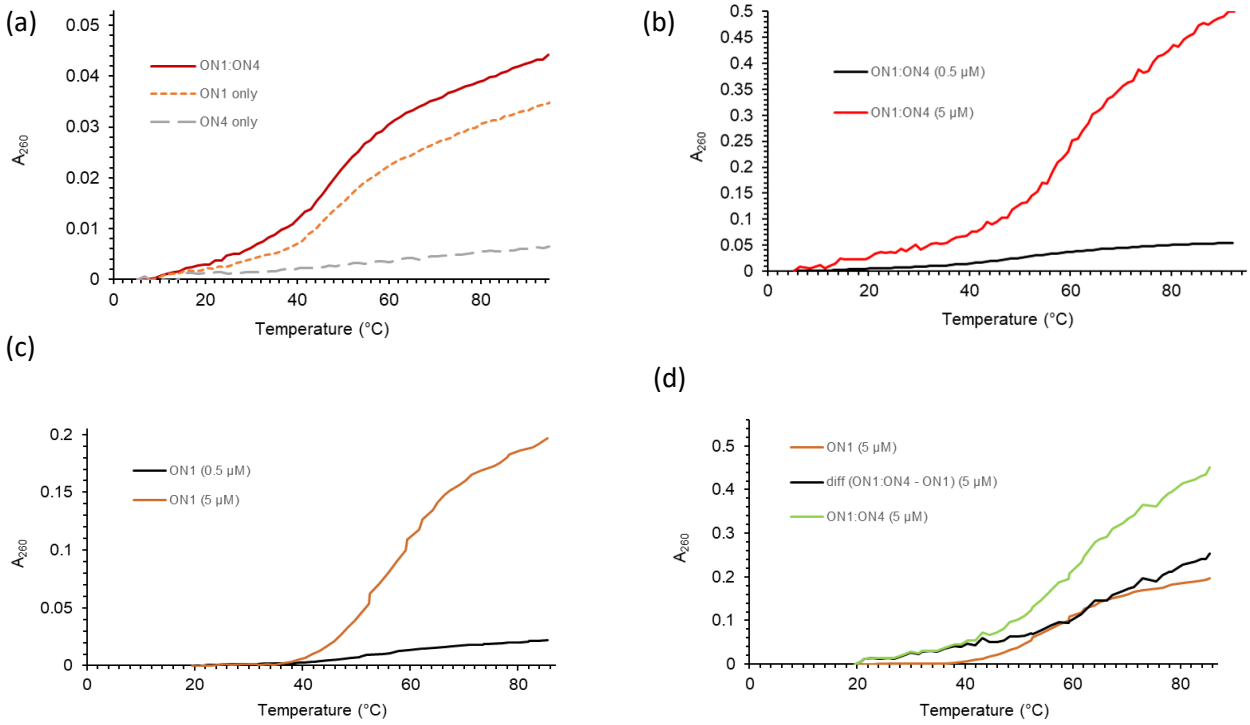


Figure S26. (a) Representative thermal denaturation profiles of **ON1**-only, **ON4**-only, and **ON1**:**ON4** in medium salt buffer. Experiments were performed using a 0.5 µM concentration of each strand. b) Thermal denaturation profiles of **ON1**:**ON4** or c) **ON1** performed using 0.5 µM or 5 µM concentration of each strand. (d) Thermal denaturation profiles of **ON1**, **ON1**:**ON4** and differential profile (**ON1**:**ON4** – **ON1**) at 5 µM concentration of each strand. For experimental conditions, see Table 1.

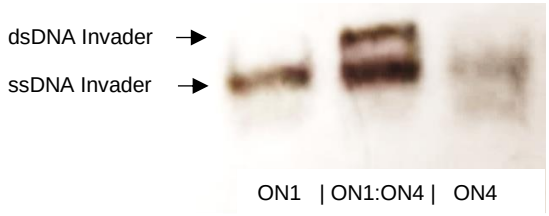


Figure S27. Representative gel electrophoretograms from non-denaturing PAGE aiming to determine if **ON1** forms a duplex with **ON4** in HEPES buffer. Experimental conditions are as outlined in Fig. 3.

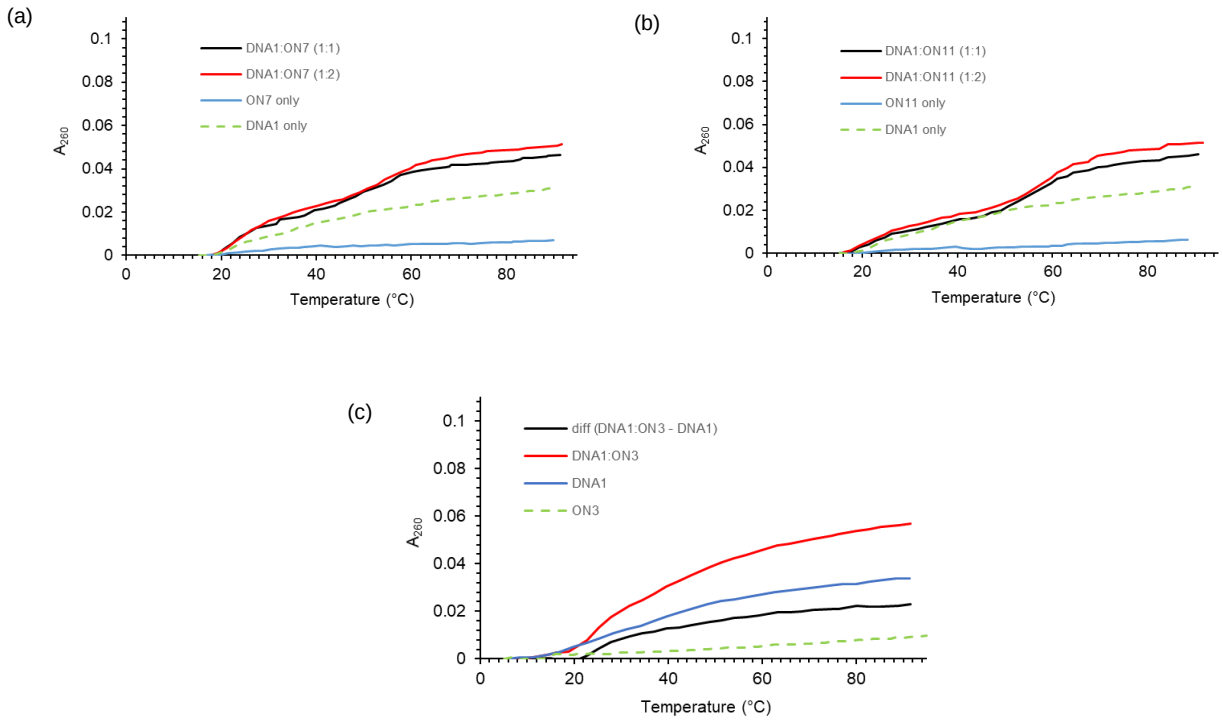


Figure S28. Additional experiments to further clarify duplex formation between **DNA1** and **ON7**, **ON11** or **ON3**. Representative thermal denaturation profiles of a) **DNA1**-only, **ON7**-only, **DNA1**:**ON7** (1:1 ratio) and **DNA1**:**ON7** (1:2 ratio), b) **DNA1**-only, **ON11**-only, **DNA1**:**ON11** (1:1 ratio) and **DNA1**:**ON11** (1:2 ratio), and c) **DNA1**-only, **ON3**-only, **DNA1**:**ON3** and differential profile (**DNA1**:**ON3** – **DNA1**-only). For experimental conditions, see Table 1.

Table S2. Sequences of DNA hairpins used in this study.[a]

DNA hairpin target	Sequence
DH1	5'- AAG CTG CAC AGG TAT ATA TAG GCC GCA TAT GCA 3'- TTC GAC GTG TCC ATA TAT ATC CGG CGT ATA CGT T_{10}
MM1	5'- AAG CTG CAC AGG TAT TTA TAG GCC GCA TAT GCA 3'- TTC GAC GTG TCC ATA AAT ATC CGG CGATATA CGT T_{10}
MM2	5'- AAG CTG GAC AGG TAT ATA TAG GCC GCT TAT GCA 3'- TTC GAC CTG TCC ATA TAT ATC CGG CGA ATA CGT T_{10}
MM3	5'- AAG CTG GAC AGG TAT TTA TAG GCC GCT TAT GCA 3'- TTC GAC CTG TCC ATA AAT ATC CGG CGA ATA CGT T_{10}
DH2	5'- ACT GTG TGT TAT ATG CTG TTC TCA GCC CTA CTG 3'- TGA CAC ACA ATA TAC GAC AAG AGT CGG GAT GAC T_{10}
DH2-MM	5'- ACT GTG TGA TAT ATG GTG TTC TCA GGC CTA CTG 3'- TGA CAC ACT ATA TAC CAC AAG AGT CCG GAT GAC T_{10}

[a] Yellow highlights indicate the position of sequence differences relative to the utilized probes.

Table S3. Quantification of **DH1**-recognition when using a 50-fold molar excess of different Invader probes.[a]

Probe	Construct	Sequence	Recognition
NIP1	6-13-6 nicked Invader	5'-UGC ACA GGT AUA UAT AGG CCG CAU A 3'-ACG TGU CCA TAU AUA TCC GGC GTA U	~30%
NIP2	8-9-8 nicked Invader	5'-UGC ACA GGT AUA UAT AGG CCG CAU A 3'-ACG TGU CCA TAU AUA TCC GGC GTA U	~40%
NIP3	10-5-10 nicked Invader	5'-UGC ACA GGT AUA UAT AGG CCG CAU A 3'-ACG TGU CCA TAU AUA TCC GGC GTA U	~25%
NIP4	12-1-12 nicked Invader	5'-UGC ACA GGT AUA UAT AGG CCG CAU A 3'-ACG TGU CCA TAU AUA TCC GGC GTA U	~15%
TIP1	6-13-6 toe. Invader	5'-UGC ACA GGT AUA UAT AGG C 3'-CCA TAU AUA TCC GGC GTA U	~20%
TIP2	8-9-8 toe. Invader	5'-UGC ACA GGT AUA UAT AG 3'-A TAU AUA TCC GGC GTA U	<15%
TIP3	10-5-10 toe. Invader	5'-UGC ACA GGT AUA UAT 3'-AU AUA TCC GGC GTA U	<5%
TIP4	12-1-12 toe. Invader	5'-UGC ACA GGT AUA U 3'-AUA TCC GGC GTA U	<5%
ON19:**ON20**	13-mer conv. Invader	5'-GGT AUA UAT AGG C 3'-CCA TAU AUA TCC G	<5%
ON17:**ON4**	19-mer conv. Invader	5'- GGT AUA UAT AGG CCG CAU A 3'- CCA TAU AUA TCC GGC GTA U	<15%
ON1:**ON18**	19-mer conv. Invader	5'-UGC ACA GGT AUA UAT AGG C 3'-ACG TGU CCA TAU AUA TCC G	<15 %

[a]Experiments were performed in triplicate. Corresponding representative electrophoretograms are shown in Fig. 3.

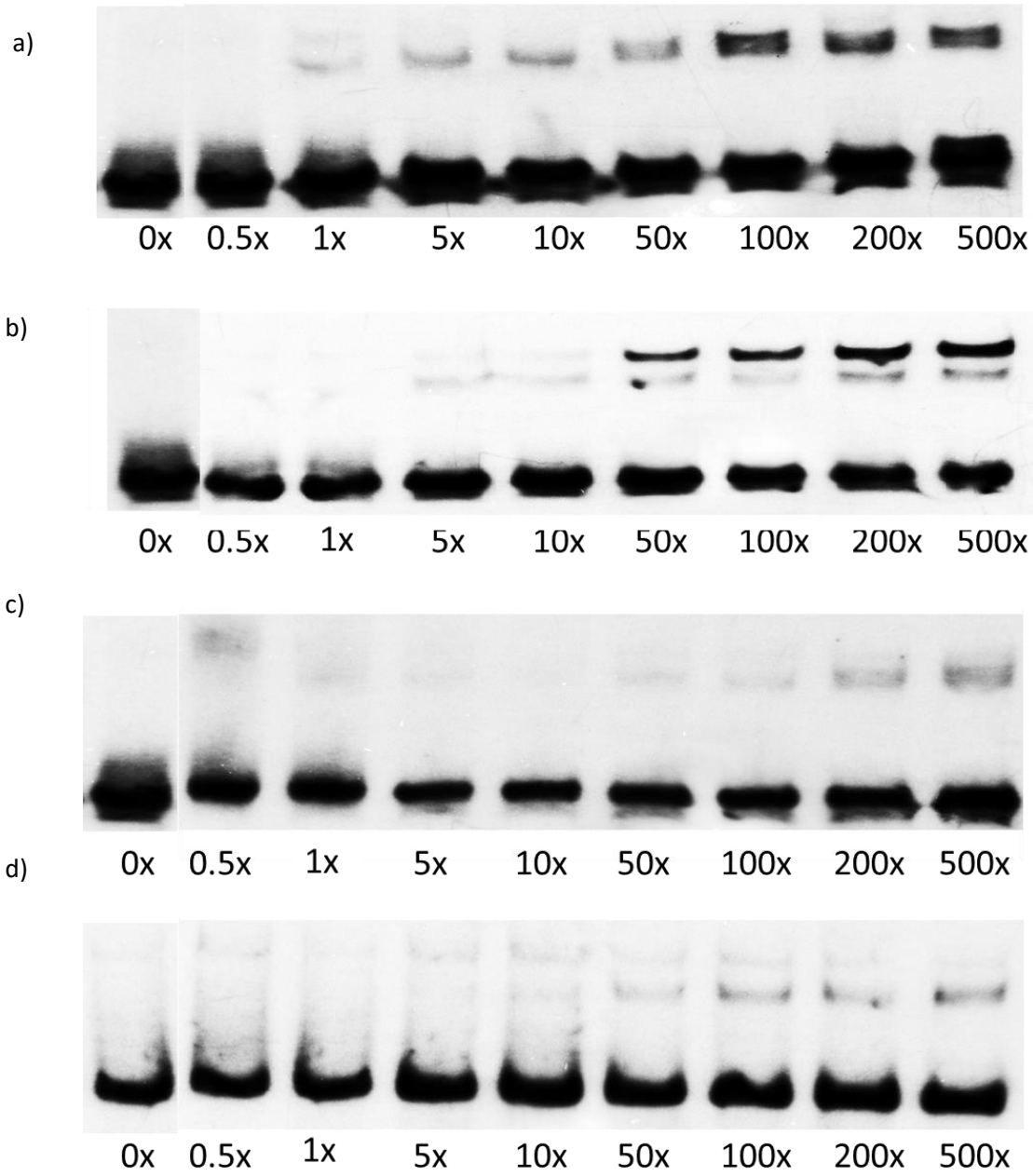


Figure S29. Representative gel electrophoretograms from dose-response experiments in which DNA hairpin **DH1** (50 nmol) was incubated with a variable molar excess of different nicked Invader probes and the corresponding toehold probes: a) **NIP1** (**ON1/2**:**ON3/4**), b) **NIP3** (**ON9/10**:**ON11/12**), c) **TIP1** (**ON1**:**ON4**) and d) **TIP3** (**ON9**:**ON12**) at 37 °C. Experimental conditions are as described in Fig. 3.

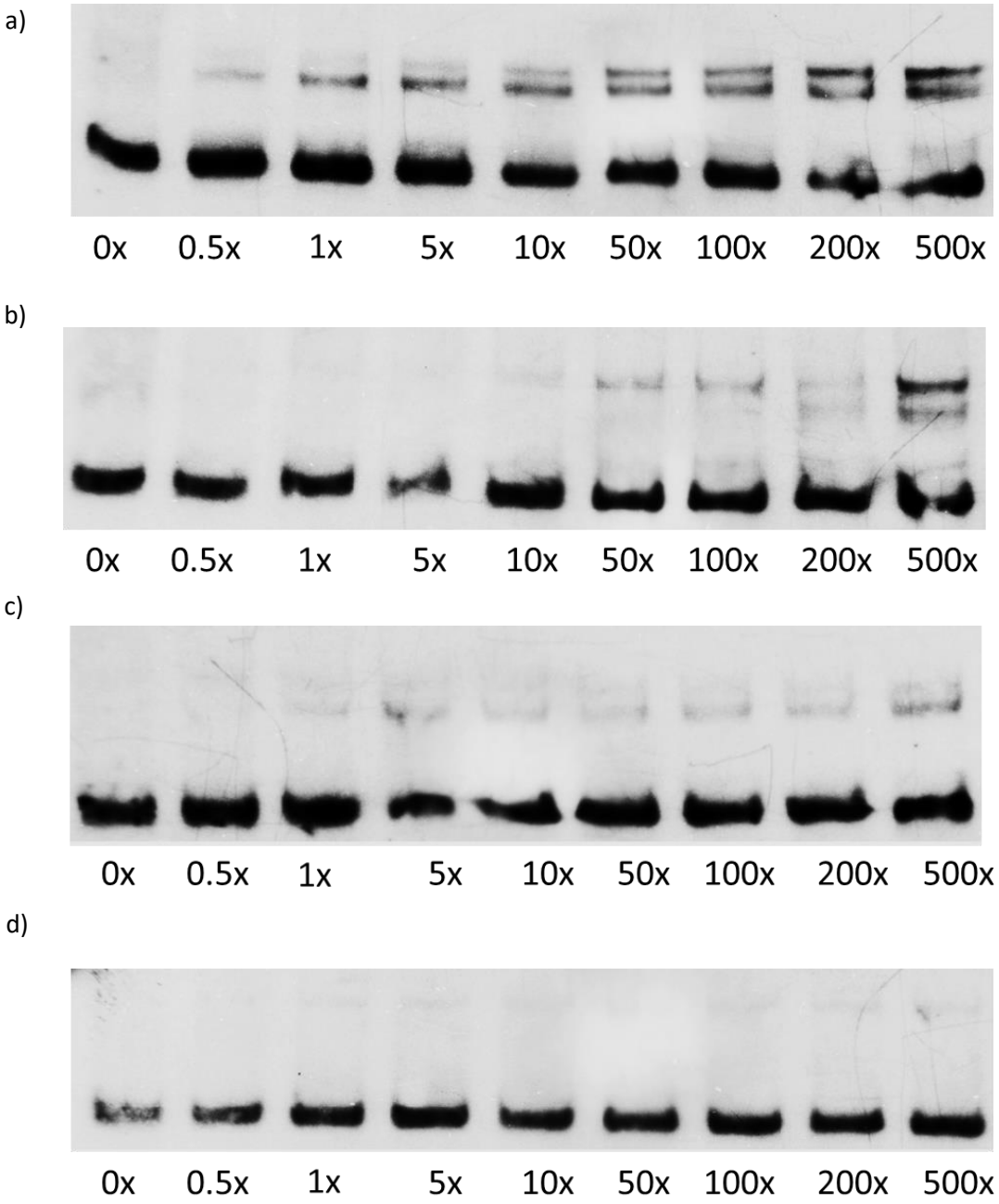


Figure S30. Representative gel electrophoretograms from dose-response experiments in which DNA hairpin **DH1** (50 nmol) was incubated with a variable molar excess of different nicked Invader probes and the corresponding toehold probes: a) **NIP2** (**ON5/6**:**ON7/8**), b) **NIP4** (**ON13/14**:**ON15/16**), c) **TIP2** (**ON5**:**ON8**), and d) **TIP4** (**ON13**:**ON16**) at 37 °C. Experimental conditions are as described in Fig. 3.

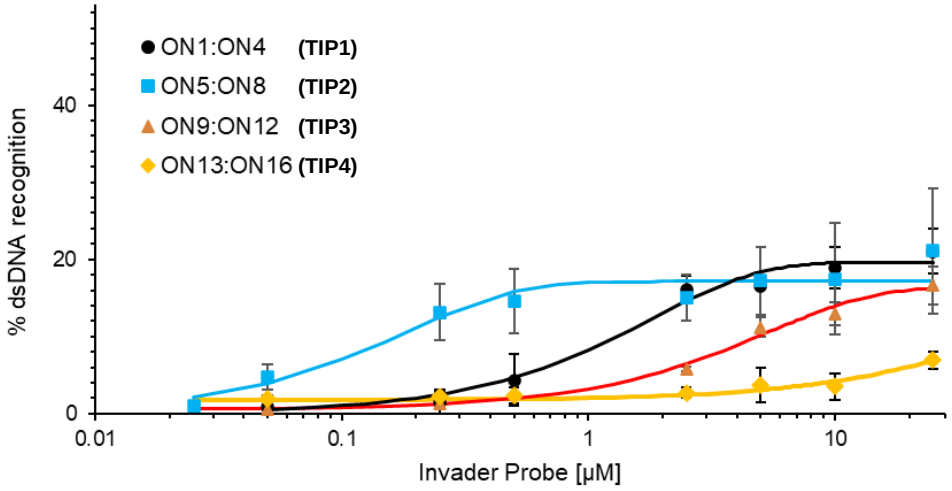


Figure S31. Dose-response profiles for recognition of **DH1** by toehold Invader probes. Curves are constructed based on the gel electrophoretograms shown in Figs. S29 and S30. Experimental conditions are as described in Fig. 3.

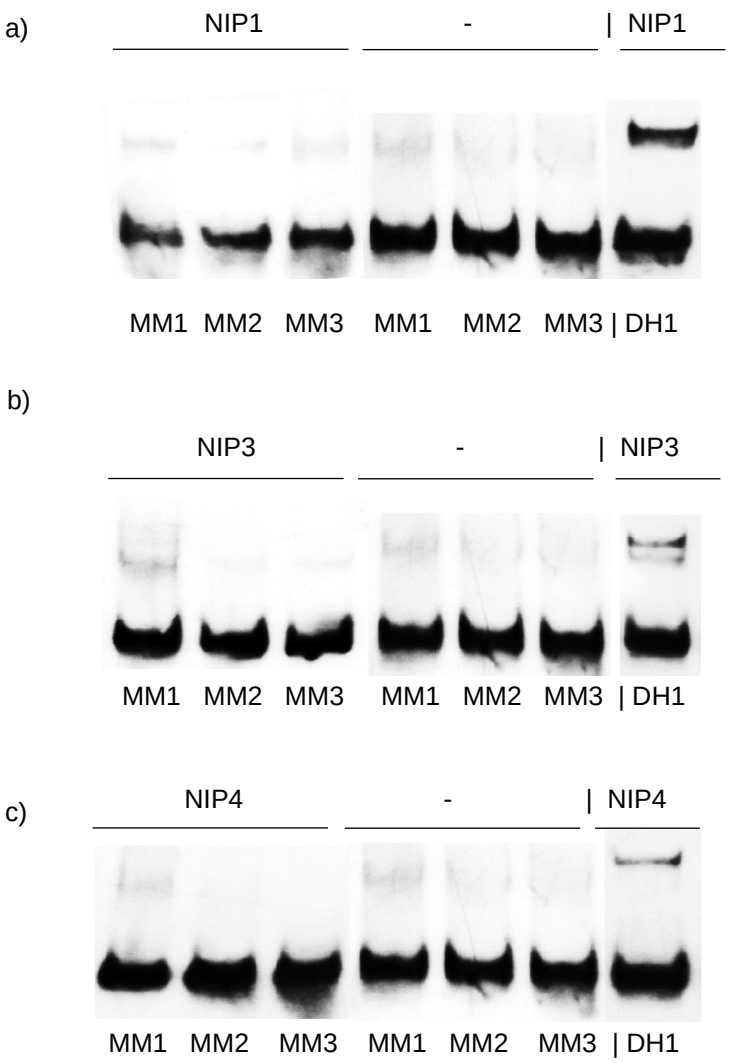


Figure S32. Binding specificities of nicked Invader probes. Representative electrophoretograms from experiments in which a 50-fold molar excess of pre-annealed a) **NIP1**, b) **NIP3** or c) **NIP4** were incubated with pre-annealed non-complementary **MM1**-**MM3** targets in HEPES buffer at 37 °C for 17 h as described in Fig. 5. For sequences of **MM1**-**MM3**, see Table S2. For an illustration of the mismatched recognition complexes that would ensue upon recognition, see Fig. S33.

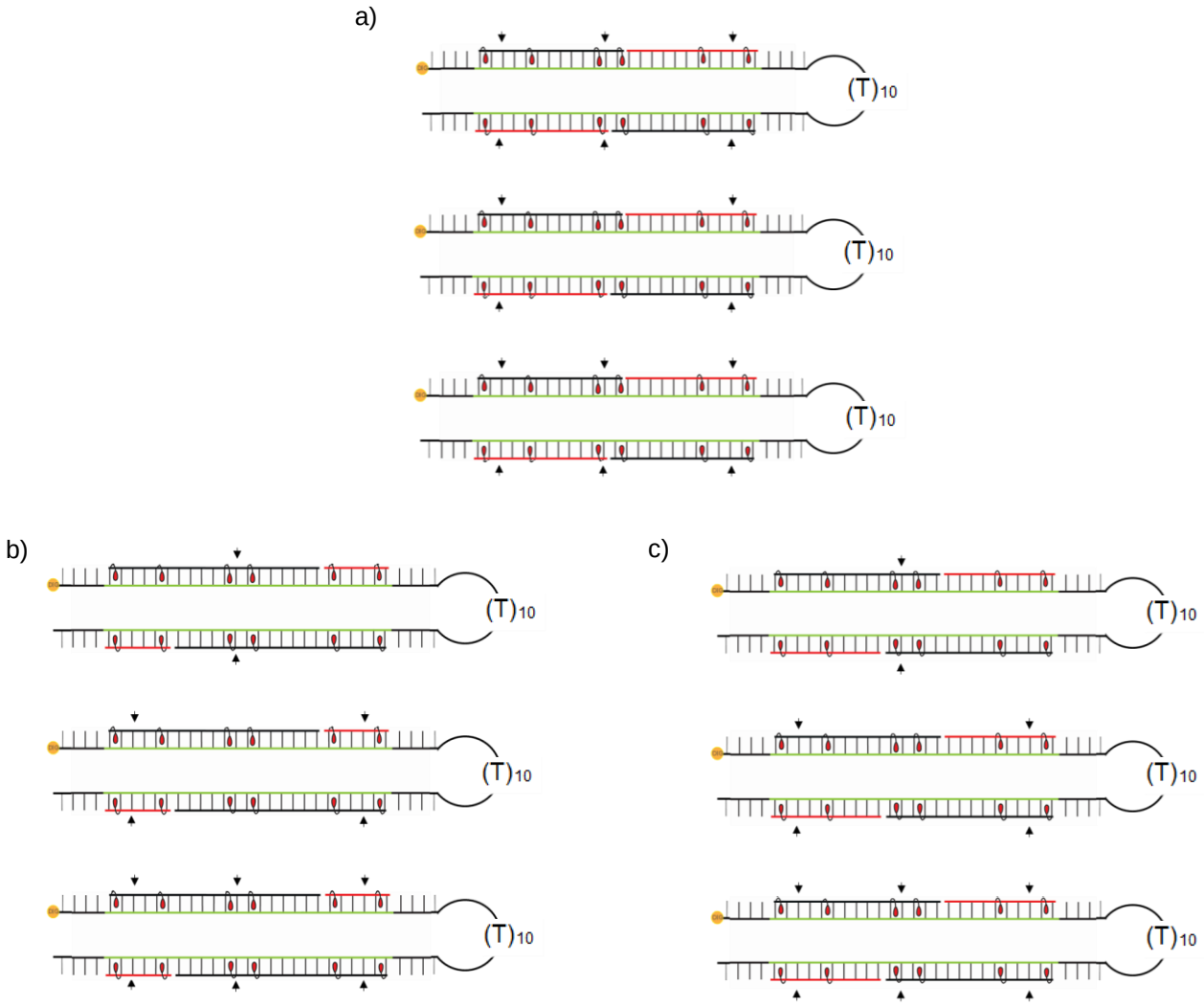


Figure S33. Illustration of the mismatched recognition complexes that would ensue upon recognition of **MM1-MM3** by a) **NIP1**, b) **NIP3**, and c) **NIP4**; arrows indicate position of mismatched base-pairs. For sequences of **MM1-MM3**, see Table S2†.

Table S4. Sequences of probes used in FISH experiments, and T_ms of probe duplexes and duplexes with DNA targets at high salt conditions.[a]

ONs	Sequence	T_m (°C)						
		5'-main probe vs 3'-aux probe	**5'-main probe vs 3'-main probe**	**5'-aux probe vs 3'-main probe**	**5'-main probe vs DNA4**	**5'-aux probe vs DNA4**	**3'-aux probe vs DNA3**	**3'-main probe vs DNA3**
23/24 **25/26**	5'-Cy3-TG**U**GT**U**ATA**U**GC**U**GTTCT**C**AGCC**C**T 3'-ACA**C**AA**U**ATA**C**GA**C**AAGAG**U**CGGG**A**-Cy3	37.0	36.0	nt	75.0	58.0	43.0	>76.0

[a] The unmodified **DNA3**:**DNA4** duplex (T_m = 82.0 °C) is the model dsDNA target, where **DNA3** = 5'-ACTGTGTGTTATATGCTGTTCTCAGCCCTACTG and **DNA4** = 3'-TGACACACAATATACGACAAGAGTCGGGATGAC. Thermal denaturation curves were recorded as described in Table 1 with the exception that in a high salt phosphate buffer was used ($[Na^+]$ = 710 mM, $[Cl^-]$ = 710 mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4), [EDTA] = 0.2 mM), with each ON present at 0.5 μM concentration. nt = no sigmoidal transition observed.

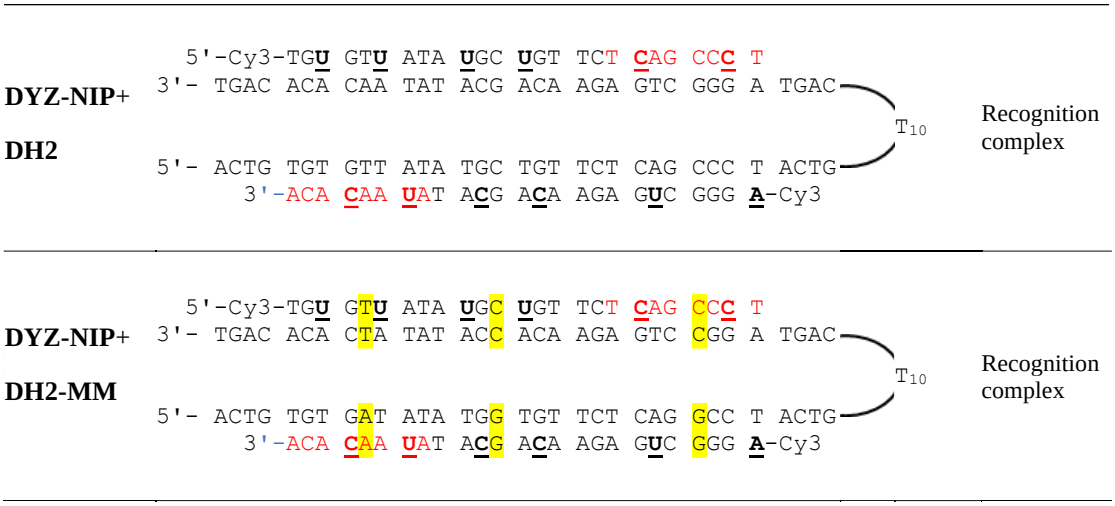


Figure S34. Matched/mismatched recognition complexes that would ensue upon recognition of complementary target **DH2** or non-complementary target **DH2-MM** by **DYZ-NIP**. Higlighted residues indicate the position of mismatched base-pairs.

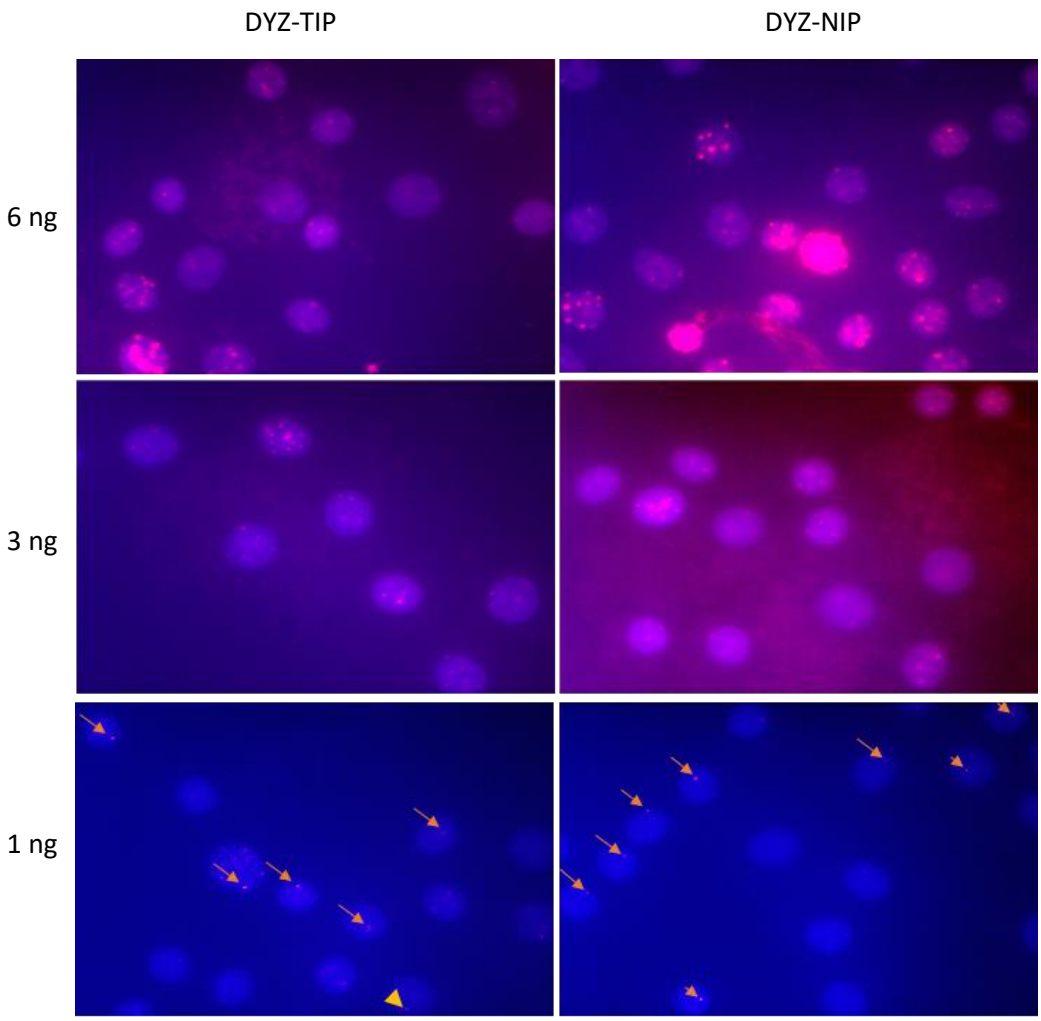


Figure S35. Images from nd-FISH experiments in which different concentrations of toehold Invader probe **DYZ-TIP** or nicked Invader probe **DYZ-NIP** (1-6 ng per 200 ul of PCR buffer) were incubated with fixed isolated nuclei from male bovine kidney cells for 3 h at 37.5 °C in a Tris buffer (20 mM Tris-Cl, 100 mM KCl, pH 8.0) and counterstained with DAPI as outlined in Fig. 7. High background was observed when **DYZ-NIP** and **DYZ-TIP** were used at 3 or 6 ng per 200 ul of PCR buffer, whereas 1 ng per 200 ul of PCR buffer yielded single Cy3-signals.

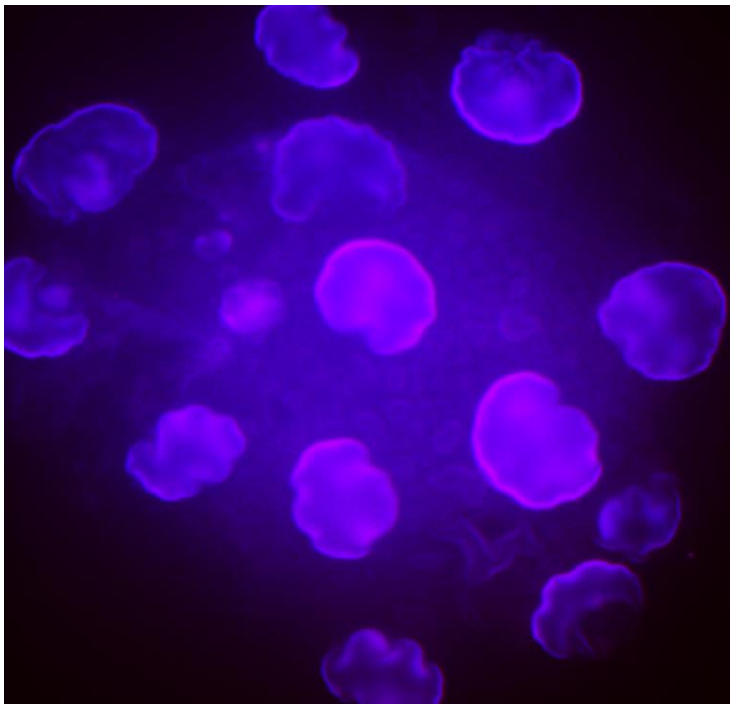


Figure S36. Image from nd-FISH experiments in which nicked Invader probe **DYZ-NIP** was incubated with isolated fixed nuclei from a female bovine endothelial cell line (15 ng per 200 µl of PCR buffer, 3 h, 37.5 °C). Experimental conditions and image analysis was carried out as described in Fig. 7.

Supplementary references

S1) S. P. Adhikari, P. Vukelich, D. C. Guenther, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2021, **19**, 9276 – 9290.

S2) C. P. Shepard, M.S thesis, University of Idaho, 2020.