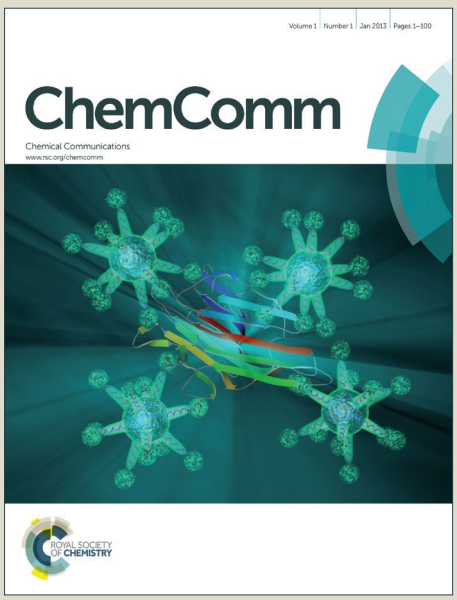
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Bulged Invader probes: Activated duplexes for mixed-sequence dsDNA recognition with improved thermodynamic and kinetic profiles†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Dale C. Guenther, Saswata Karmakar and Patrick J. Hrdlicka*

Double-stranded oligonucleotides with +1 interstrand zipper arrangements of intercalator-functionalized nucleotides are energetically activated for recognition of mixed-sequence double-stranded DNA. Incorporation of nonyl (C_9) bulges at specific positions of these probes, results in more highly affine (>5-fold), faster (>4-fold) and more persistent dsDNA recognition relative to conventional Invader probes.

Chemical probes capable of sequence-specific recognition of dsDNA have tremendous potential as tools in diagnostics, structural elucidations, and nanotechnology.[1-5] Hybridization-based approaches are particularly interesting due to their predictable binding modes and the resulting ease of design. To realize sequence-specific dsDNA recognition, probes must invade Watson-Crick base pairs or bind via extrahelical contacts such as Hoogsteen base-pairing, with triplex-forming oligonucleotides[1,6] and peptide nucleic acids (PNAs)[4,7] as prime examples of the latter. However, triplex-based approaches rely on the presence of long polypurine regions, which limits the number of targetable sites. In contrast, conformationally restricted γ-PNAs[8] bind to complementary DNA (cDNA) with sufficient affinity to invade Watson-Crick base-pairs of dsDNA targets, albeit only at non-physiologic ionic strengths, resulting in displacement of one target strand and formation of a D-loop.

Double-stranded probes that bind to dsDNA via double-duplex invasion, offer the promise of even more favorable binding thermodynamics and improved specificity, as binding to mismatched dsDNA regions generates two destabilized duplexes.[9] However, the probe duplex must dissociate easily for this approach to be effective. One strategy to realize this has been through the use of pseudocomplementary (pc) base pairs such as 2,6-diaminopurine and 2-thiouracil, which form weak base-pairs with each other, while forming stable pairs with thymine and adenine in target strands.[10] The energy difference between the double-stranded probe and the resulting probe-target duplexes generates a thermodynamic gradient for dsDNA recognition. While pcDNA only are weakly activated for dsDNA recognition,[11] pcPNA have been shown to recognize internal regions of mixed-sequence dsDNA at low ionic strengths.[12]

As part of our efforts toward developing new strategies for mixed-sequence dsDNA recognition, we recently introduced so-called Invader probes, which also rely on energy differences between probe duplexes and recognition complexes to drive dsDNA recognition (Figure 1).[13] These probes feature 2'-intercalator-functionalized nucleotides that are arranged in +1 interstrand zipper motifs, which force the covalently linked intercalators to compete for the same inter-base-pair region, leading to violation of the nearest-neighbor exclusion principle[14] and probe destabilization.[13,15-19] In the recognition complex, in which each probe strand is bound to a complementary DNA region, the intercalators no longer compete for the same space, leading to strong duplex stabilization due to efficient π-π-stacking interactions with neighboring base-pairs. In previous studies, we have: i) identified more easily accessible analogs of the N2'-pyrene-functionalized 2'-amino-α-L-LNA (Locked Nucleic Acid) monomers that were used in original Invader designs,[15] which include the 2'-*O*-(pyren-1-yl)methyl-RNA monomer shown in Figure 1, ii) studied the influence that the intercalator, linker, nucleobase, and number and distance between the intercalator-functionalized nucleotides[13,15-19] have on dsDNA recognition efficiency, and iii) demonstrated recognition of chromosomal DNA targets at non-denaturing conditions.[19]

Herein, we describe improved dsDNA recognition using a novel Invader probe architecture that contains non-nucleosidic nonyl (C_9) bulge inserts (Figure 1). This design was pursued based on the hypothesis that internal C_9 bulges would destabilize the probe duplex, promote local denaturation, thus revealing the Watson-Crick face of the probe, and accelerate nucleation with, and invasion of, dsDNA targets.

*Department of Chemistry, University of Idaho, 875 Perimeter Dr, Moscow, ID 83844-2343, USA. E-mail: hrdlicka@uidaho.edu

† Electronic Supplementary Information (ESI) available: Experimental protocols; MS data for modified ONs; representative thermal denaturation curves; additional gel electrophoretograms, kinetics plots, and T_m and dsDNA-recognition data. See DOI: 10.1039/x0xx00000x

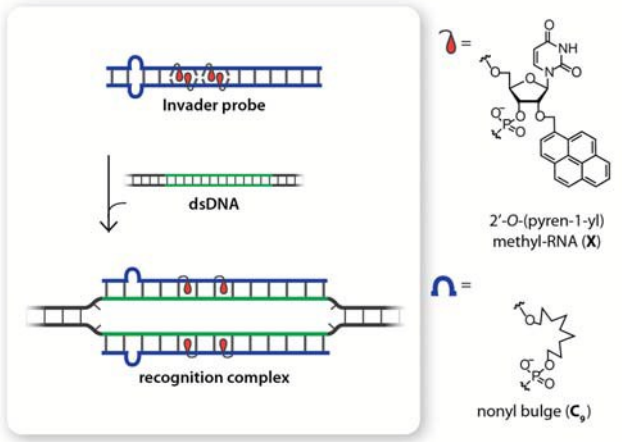


Figure 1. Schematic representation of dsDNA recognition by Invader probes containing non-nucleosidic bulges and the chemical modifications used for this approach.

Bulges have been used to tune the hybridization properties of oligonucleotides.[20,21] While they induce minimal perturbation of the global duplex conformation, they do destabilize duplexes by interrupting the π-stack.[21] By adjusting the number and position of the C_9 bulges, we hypothesized that we could destabilize probe duplexes more than probe-target duplexes, resulting in a more prominent thermodynamic driving force and faster dsDNA recognition.

A library of Invader probes, containing two consecutive +1 interstrand zipper motifs of 2′-*O*-(pyren-1-yl)methyl-RNA-U monomers at the center and one or two C_9 bulges at one or both termini, were synthesized (Table 1). Thermal denaturation temperatures (T_m's) of these probes and the duplexes with cDNA were compared to conventional Invaders without C_9 bulges. As expected from our previous work, reference Invader strands **ON1** and **ON2** form very stable duplexes with cDNA (ΔT_m = 18 °C relative to unmodified ON).[19] The insertion of a single C_9 bulge into an Invader strand greatly reduces T_m's (-9 to -12 °C) relative to **ON1** or **ON2**. Insertion of two C_9 bulges potentiates these trends (T_m < 15 °C for **ON7** or **ON8** vs cDNA). The double-stranded Invader probes display significantly lower T_m's than the corresponding duplexes between individual probe strands and cDNA, verifying our previous observations that +1 interstrand zipper motifs of **X** monomers are inherently destabilizing (e.g., compare T_m of **ON1**:**ON2** vs **ON1**:cDNA and **ON2**:cDNA). Invader probes, in which two C_9 bulges either are present on the same strand or on two different strands but the same terminus, are particularly destabilized.

The thermodynamic dsDNA recognition potential of a specific Invader probe can be estimated by the term *thermal advantage*, given as *TA* = T_m (5′-Inv:cDNA) + T_m (3′-Inv:cDNA) - T_m (Invader probe) - T_m (dsDNA target), with large positive values signifying a strongly activated probe. Invader probe **ON1**:**ON2**, which is based on a traditional probe architecture without bulges, has a prominent *TA* value of 28.5 °C due to the high T_m's of probe:cDNA duplexes and low T_m of the probe duplex.

Table 1. Thermal denaturation temperatures (T_m's) and thermal advantages (*TA*'s) for modified DNA duplexes.[a]

Probe	Sequence	5'-Inv: 3'-Inv	T_m [ΔT_m] (°C) 5′-Inv: cDNA	T_m [ΔT_m] (°C) 3′-Inv: cDNA	*TA* (°C)
1:2	5′-GGTAXIAXIATAGGC 3′-CCATAIXAIXATCCG	45.0 [+7.5]	55.5 [+18.0]	55.5 [+18.0]	28.5
3:2	5′-GG TAXIAXIATAGGC 3′-CC ATAIXAIXATCCG	31.5 [-6.0]	44.0 [+6.5]	55.5 [+18.0]	30.5
5:2	5′-GGTAXIAXIATAG GC 3′-CCATAIXAIXATC CG	33.0 [-4.5]	44.5 [+7.0]	55.5 [+18.0]	29.5
1:4	5′-GG TAXIAXIATAGGC 3′-CC ATAIXAIXATCCG	35.0 [-2.5]	55.5 [+18.0]	46.5 [+9.0]	29.5
1:6	5′-GGTAXIAXIATAG GC 3′-CCATAIXAIXATC CG	28.5 [-9.0]	55.5 [+18.0]	43.5 [+6.0]	33.0
3:4	5′-GG TAXIAXIATAGGC 3′-CC ATAIXAIXATCCG	<15.0	44.0 [+6.5]	46.5 [+9.0]	>38.0
5:6	5′-GGTAXIAXIATAG GC 3′-CCATAIXAIXATC CG	<15.0	44.5 [+7.0]	43.5 [+6.0]	>35.5
7:2	5′-GG TAXIAXIATAG GC 3′-CC ATAIXAIXATC CG	<15.0	<15.0	55.5 [+18.0]	-
1:8	5′-GG TAXIAXIATAG GC 3′-CC ATAIXAIXATC CG	<15.0	55.5 [+18.0]	<15.0	-
3:6	5′-GG TAXIAXIATAG GC 3′-CC ATAIXAIXATC CG	28.5 [-9.0]	44.0 [+6.5]	43.5 [+6.0]	21.5
5:4	5′-GG TAXIAXIATAG GC 3′-CC ATAIXAIXATC CG	32.5 [-5.0]	44.5 [+7.0]	46.5 [+9.0]	21.0

[a] ΔT_m is calculated relative to the corresponding unmodified dsDNA (T_m = 37.5 °C; Thermal denaturation curves were recorded in medium salt phosphate buffer ([Na^+] = 110 mM, [Cl^-] = 100 mM, pH 7.0 (NaH_2PO_4/Na_2HPO_4), [EDTA] = 0.2 mM) and each [ON] = 0.5 μM; see main text for definition of *TA*.

Invader probes with a single C_9 bulge (e.g., **ON3**:**ON2**) display similar or slightly higher *TA*s since the bulge destabilizes probe:cDNA and Invader probe duplexes to similar degrees. Probes **ON3**:**ON4** and **ON5**:**ON6**, which have two C_9 bulges at one of the termini, display significantly increased dsDNA recognition potential (*TA*s > 35.5 °C), because the probe duplexes are very strongly destabilized, while the probe-target duplexes only are mildly destabilized; presumably, this is because two adjacent C_9 bulges (as in probe duplexes) have a more detrimental effect on base-pairing cooperativity than two separate C_9 bulges (as in probe-target duplexes). In line with this, Invader probes with two C_9 bulges on separate strands and termini (**ON3**:**ON6** and **ON5**:**ON4**) display lower dsDNA recognition potential because the probe duplexes are not as destabilized. *TA* values for Invader probes with two C_9 bulges on one strand (**ON7**:**ON2** and **ON1**:**ON8**) could not be determined due to the low stability of probe-target duplexes.

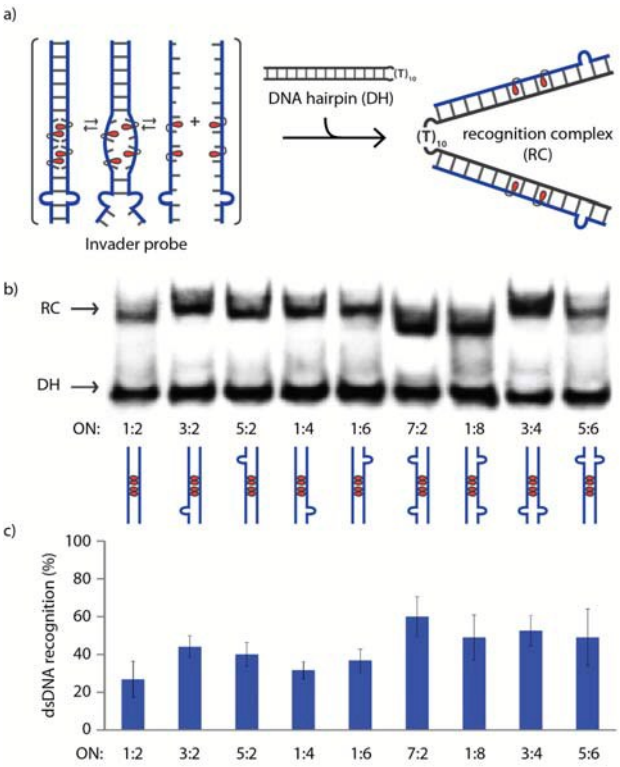


Figure 2. (a) Schematic representation of the EMSA used to evaluate dsDNA recognition of Invader probes. (b) Representative electrophoretograms for recognition of model dsDNA target **DH1** (34.4 μM) by different Invader probes (6.88 μM) at 8 °C. (c) Histogram showing the average of three experiments; error bars represent standard deviation. DIG-labeled **DH1** (5′-GGTATATATAGGC-T_{10}-GCCTATATATACC-3′) was incubated with pre-annealed Invader probe in HEPES buffer (50 mM HEPES, 100 mM NaCl, 5 mM $MgCl_2$, pH 7.2, 10% sucrose, 1.44 mM spermine tetrahyrdochloride) for 17 h.

TA values provide an estimate for the thermodynamic dsDNA recognition potential of specific Invader probes.‡ However, other factors, including the experimental temperatures used, likely influence recognition efficiency and kinetics. To elucidate this, an electrophoretic mobility shift assay (EMSA) was performed. Pre-annealed Invader probes were incubated with DNA hairpin **DH1**, in which the double-stranded target region is linked via a decameric thymidine loop (Figure 2a). Recognition of this model target results in the formation of a recognition complex, which is observed as a slower moving band on non-denaturing polyacrylamide gel electrophoresis (Figure 2b). A 200-fold molar excess of Invader probes was incubated with **DH1** at 8 °C for 17 h. At these conditions, the conventional Invader probe **ON1**:**ON2** only results in ~22% recognition, whereas single bulge Invaders result in more efficient recognition (30-42%) (Figure 2c and Table S2†). Invader probes with two C_9 bulges at one terminus (**ON3**:**ON4** and **ON5**:**ON6**) or two C9 bulges on the same strand (**ON1**:**ON8** and **ON7**:**ON2**) recognize the dsDNA target even more efficiently (41-55%). The recognition complexes formed with **ON1**:**ON8** and **ON7**:**ON2** have slightly greater electrophoretic mobilities than those formed with other Invader probes. This is almost certainly because binary, rather than ternary, recognition complexes are formed, as **ON7** and **ON8** have very low cDNA affinity (T_m < 15 °C for **ON7**/**ON8**:cDNA, Table 1 – see also Figure S2†). Invader

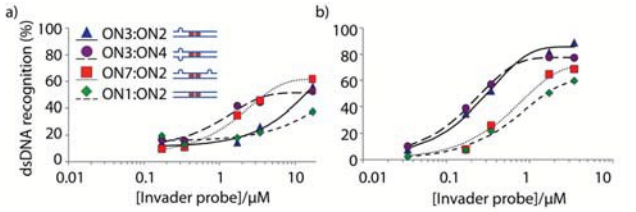


Figure 3. Dose-response curves for recognition of dsDNA by Invader probes **ON3**:**ON2**, **ON3**:**ON4**, **ON7**:**ON2**, and **ON1**:**ON2** at (a) 8 °C or (b) 22 °C. Experimental conditions otherwise as described in Figure 2.

probes with two C_9 bulges on separate strands and termini (**ON3**:**ON6** and **ON5**:**ON4**) do not result in detectable dsDNA recognition, suggesting that the process is energetically unfavorable (Figure S3†). For similar reasons, Invader probes with three or four bulge insertions also do not result in detectable dsDNA recognition (Figure S3 and Table S3†).

While conventional Invader strands **ON1** and **ON2** result in some recognition of **DH1** when used as single-stranded probes, none of the C_9-containing single-stranded probes result in significant recognition of **DH1** (Figure S4†). Interestingly, **ON7**:**ON2** results in more pronounced dsDNA recognition than single-stranded **ON2**, indicating that the presence of **ON7** is advantageous despite its low cDNA affinity (Figure S2†).

Dose-response assays were performed at 8 °C or ambient temperature (22 °C) for representative Invader probes (Figure 3). At ambient temperature, single bulge Invader **ON3**:**ON2** and Invader **ON3**:**ON4**, which has two bulges at the same terminus, display similar dose-response profiles and sub-micromolar C_{50} values (i.e. the probe concentration resulting in 50% recognition of **DH1**; ~0.3 μM, Table 2). Conventional Invader probe **ON1**:**ON2** has a significantly higher C_{50} value (~1.6 μM), whereas Invader **ON7**:**ON2**, with two bulges on the same strand, has an intermediate C_{50} value of ~1.0 μM. Incubation at 8 °C results in slightly different dose-response trends (compare Figures 3a and 3b). Thus, double bulge Invaders **ON3**:**ON4** and **ON7**:**ON2** display lower C_{30} values, than single bulge Invader **ON3**:**ON2** or conventional Invader **ON1**:**ON2**. These observations suggest that probes with large thermodynamic driving forces result in more efficient dsDNA recognition at higher experimental temperatures, whereas probes with low T_m's result in efficient dsDNA recognition at low experimental temperatures where breathing of base-pairs is minimal. Probes with low T_m's are likely partially or even fully dissociated at low experimental temperatures, thereby enabling the Watson-Crick face of the probe strands to be available for nucleation with DNA targets.

The kinetics of Invader-mediated dsDNA recognition were determined in experiments in which a 100-fold molar excess of probe was incubated with **DH1** at 22 °C (Figure 4). All of the bulge-containing Invaders display much faster recognition kinetics than conventional Invader probe **ON1**:**ON2** (pseudo-first order rate constants shown in Table 2). Invader probes **ON3**:**ON2**, **ON3**:**ON4** and **ON7**:**ON2** display 2.3, 2.7 and 4.1-fold faster kinetics, respectively. Presumably, the bulges promote partial or even full denaturation of the Invader probes, thus revealing their Watson-Crick face for faster target binding.

The persistence of dsDNA-binding was evaluated in a

Table 2. Summary of parameters for dsDNA recognition by representative Invader probes.

ON	C_{50}[a] 22 °C (μM)	C_{30}[a] 8 °C (μM)	t_{50}[b] (min)	k_{obs}[c] (10^{-3} min^{-1})	k_{rel}
1:2	1.6	9.7	-	3.1	1
3:2	0.3	5.2	110	7.2	2.3
3:4	0.3	1.0	42	8.4	2.7
7:2	1.0	1.5	41	13	4.1

[a] Calculated from curves shown in Figure 3. [b] t_{50} = time to reach 50% dsDNA recognition at 22 °C as calculated from time-course experiments shown in Figure 4. [c] Calculated from the pseudo-first order plots shown in Figure S5†.

competition assay,[19] in which pre-formed complexes (24 h incubation at 22 °C) were challenged with a 1000-fold excess of linear dsDNA target (Figure 4). Dissociating Invader strands bind to this competitor target,[13] resulting in formation of a faster moving band in non-denaturing gel electrophoresis consistent with re-formation of **DH1**. Approximately 25% of the recognition complexes between **DH1** and **ON1**:**ON2** or **ON3**:**ON2** remain intact 6 h post-challenge. The recognition complex between **DH1** and **ON3**:**ON4**, undergoes rapid dissociation (>90% within 6 h), likely due to the low cDNA affinity of **ON3** and **ON4**. Surprisingly, the recognition complex between **DH1** and **ON7**:**ON2** is remarkably stable (~60% of complex intact after 24 h). This construct is unique, as only one probe strand (i.e., **ON2**) is firmly bound to the target in the recognition complex (Figure S2†). Given the slower dissociation of **DH1**:(**ON7**):**ON2** relative to **DH1**:**ON1**:**ON2**, it is clear that the unbound **ON7** plays a role in slowing down dissociation, possibly due to transient binding to the binary complex and/or weak affinity toward the target competitor strand.

In conclusion, probes with appropriately positioned non-nucleosidic bulges display faster, more efficient, and longer-lasting recognition of mixed-sequence dsDNA targets than conventional Invader probes. The robustness and simplicity of design render these optimized probes amenable to a variety of applications in molecular diagnostics and DNA nanotechnology.

Support for this study was provided by Award GM088697 from the National Institute of General Medical Sciences (NIH), and Awards IF13-001 and IF14-012 from the Higher Education Research Council, Idaho State Board of Education

Notes and references

‡ Thermodynamic data could not be obtained via the van't Hoff method as denaturation curves lacked clear base lines.

1 R. Besch, C. Giovannangeli and K. Degitz, *Curr. Drug Targets*, 2004, **5**, 691.
2 D. Ackermann and M. Famulok, *Nucleic Acids Res.*, 2013, **41**, 4729.
3 Y. Aiba, J. Sumaoka and M. Komiyama, *Chem. Soc. Rev.*, 2011, **40**, 5657.
4 P. E. Nielsen, *Chem. Biodiv.*, 2010, **7**, 786.
M. S. Blackledge and C. Melander, *Bioorg. Med. Chem.*, 2013, **21**, 6101.
5 T. Gai, C. A. Gersbach and C. F. Barbas III, *Trends Biotechnol.*, 2013, **31**, 397.

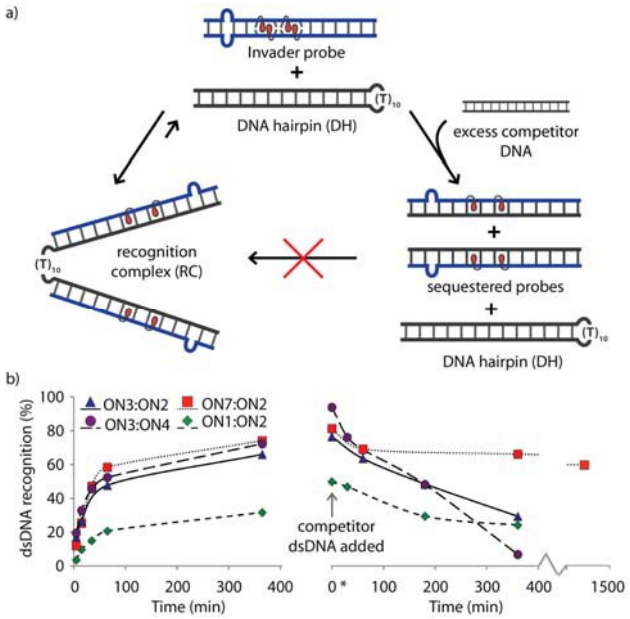


Figure 4. a) Assays used to determine kinetic parameters for dsDNA recognition using representative Invader probes. b) *Left*: Kinetics of recognition complex formation at 22 °C using 100-fold molar excess of Invader probes. *Right*: Competitive dissociation kinetics of recognition complexes between DNA hairpins and Invader probes (for representative gel electrophoretograms, see Figure S6†). 100-fold molar excess of Invader probes (3.44 μM) was incubated with **DH1** for 24 h, followed by addition of a 1000-fold molar excess of linear competitor dsDNA target (34.4 μM – sequence: 5′-GGTATATATAGGC:3′-CCATATATATCCG). *T* = 22 °C.

6 M. Duca, P. Vekhoff, K. Oussedik, L. Halby and P. B. Arimondo, *Nucleic Acids Res.*, 2008, **36**, 5123.
7 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.
8 R. Bahal, B. Sahu, S. Rapireddy, C-M. Lee and D. H. Ly, *ChemBioChem*, 2012, **13**, 56.
9 S. X. Chen, D. Y. Zhange and G. Seelig, *Nature Chem.*, 2013, **5**, 782.
10 I. V. Kutyavin, R. L. Rhinehart, E. A. Lukhtanov, V. V. Gorn, R. B. Meyer Jr. and H. B. Gamper Jr., *Biochemistry*, 1996, **35**, 11170.
11 I. V. Smolina and V. V. Demidov, *Chem. Biol.*, 2003, **10**, 591.
12 J. Lohse, O. Dahl and P. E. Nielsen, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 11804.
13 S. P. Sau, T. S. Kumar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2010, **8**, 2028.
14 D. M. Crothers, *Biopolymers*, 1968, **6**, 575.
15 S. P. Sau, A. S. Madsen, P. Podbevsek, N. K. Andersen, T. S. Kumar, S. Andersen, R. L. Rathje, B. A. Anderson, D. C. Guenther, S. Karmakar, P. Kumar, J. Plavec, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2013, **78**, 9560.
16 S. Karmakar, D. C. Guenther and P. J. Hrdlicka, *J. Org. Chem.*, 2013, **78**, 12040.
17 S. Karmakar, A. S. Madsen, D. C. Guenther, B. C. Gibbons and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2014, **12**, 7758.
18 B. A. Anderson, J. J. Onley and P. J. Hrdlicka, *J. Org. Chem.*, 2015, **80**, 5395.
19 D. C. Guenther, G. H. Anderson, S. Karmakar, B. A. Anderson, B. A. Didion, W. Guo, J. P. Verstegen and P. J. Hrdlicka, *Chem. Sci.*, 2015, **6**, 5006.
20 W. Braunlin, J. Volker, G. E. Plum and K. J. Breslauer, *Biopolymers*, 2013, **99**, 408.
21 D. V. Pyshnyi, A. A. Lomzov, I. A. Pyshnaya and E. M. Ivanova, *J. Biomol. Struct. Dyn.*, 2006, **23**, 567.