

Showcasing research from Professor Hrdlicka's laboratory,  
Department of Chemistry, University of Idaho, Moscow, USA

DNA-targeting Invader probes: discovery, principles and applications

The development of chemically modified oligonucleotides, nucleic acid mimics, and protein-based constructs capable of sequence-unrestricted recognition of specific double-stranded (ds) DNA regions continues to attract considerable attention. This Account describes the discovery, design principles, and applications of Invader probes. These double-stranded oligonucleotides incorporate intercalator-functionalized nucleotides arranged to destabilize the probe duplex, while individual strands bind complementary DNA with high affinity. This thermodynamic asymmetry enables selective recognition of dsDNA targets *via* double-duplex strand invasion, supporting their use as diagnostic agents, modulators of gene expression, and versatile tools in biotechnology.

Image reproduced by permission of Michaela E. Everly from *RSC Chem. Biol.*, 2026, **7**, 323.

As featured in:



See Patrick J. Hrdlicka and  
Michaela E. Everly,  
*RSC Chem. Biol.*, 2026, **7**, 323.



Cite this: *RSC Chem. Biol.*, 2026, 7, 323

Received 7th November 2025,  
Accepted 7th January 2026

DOI: 10.1039/d5cb00286a

rsc.li/rsc-chembio

## DNA-targeting Invader probes: discovery, principles and applications

Patrick J. Hrdlicka \* and Michaela E. Everly

Development of chemically modified oligonucleotides, nucleic acid mimics, protein-based constructs, and other ligands – capable of sequence-unrestricted recognition of specific double-stranded (ds) DNA regions – is an area of research that continues to attract considerable attention. Efforts are fueled by the need for diagnostic agents, modulators of gene expression, and novel therapeutic modalities against genetic diseases. While pioneering approaches focused on accessing nucleotide-specific features from the grooves of DNA duplexes, recent developments have entailed strand-invading probes, *i.e.*, probes capable of binding to DNA duplexes by breaking existing Watson–Crick base pairs and forming new, more stable base pairs. For the past twenty years, our laboratory has pursued the development of a type of dsDNA-targeting strand-invading probes, which we have named Invader probes. These double-stranded oligonucleotide probes feature intercalator-functionalized nucleotides that are specifically arranged to promote destabilization of the probe duplex, whereas individual strands exhibit very high affinity towards complementary DNA. This account details the discovery, principles, and applications of Invader probes.

Department of Chemistry, University of Idaho, Moscow, Idaho 83844-2343, USA. E-mail: hrdlicka@uidaho.edu



Patrick J. Hrdlicka

*Dr Hrdlicka earned his PhD in 2006 from the Nucleic Acid Center at the University of Southern Denmark, where he conducted research on functionalized locked nucleic acids (LNA) under the supervision of Prof. Jesper Wengel. He joined the Department of Chemistry at the University of Idaho later that year and was promoted to full professor in 2017. Dr Hrdlicka has received several recognitions from the University of Idaho, including the*

*Excellence in Research and Creative Activity Award in 2013. A central tenet of his academic career has been graduate mentorship, with ~15 students completing their degrees under his supervision, many of whom now work at leading oligonucleotide companies. His research group focuses on the synthesis and biophysical characterization of chemically modified oligonucleotides, with a strong emphasis on the development of strategies for sequence-unrestricted targeting of chromosomal DNA.*



Michaela E. Everly

*Dr Everly received her PhD in Chemistry in 2025 from the University of Idaho under the supervision of Dr Patrick J. Hrdlicka. Her work focused on the synthesis and characterization of novel oligonucleotide probes for highly specific and efficient recognition of mixed-sequence double-stranded DNA. She was recognized with the University of Idaho College of Science Dean's Award for Most Outstanding Graduate Student*

*in the Department of Chemistry. She is currently a postdoctoral research fellow at the University of Michigan, where she assists in the development and operation of the Nucleic Acid Therapeutics Core and investigates antisense oligonucleotide therapeutics under the mentorship of Dr Michelle Hastings.*



## Introduction

Over the past several decades, considerable efforts have been devoted to the development of synthetic probes capable of recognizing specific chromosomal DNA regions. These efforts have been fueled by the prospect of diagnostic agents, modulators of gene expression, and novel therapeutic modalities against genetic diseases. However, while more than twenty RNA-targeting antisense oligonucleotides (ONs), splice-switching ONs, and small interfering RNAs (siRNAs) have received regulatory approval in recent years as therapeutics against genetic diseases,<sup>1</sup> DNA-targeting ONs have yet to achieve similar success. This reflects the additional challenges that DNA targeting poses, such as the need for nuclear delivery of DNA-targeting ONs and high compaction of chromosomal DNA. Moreover, the DNA duplex only offers few nucleotide-specific handles that can be accessed by synthetic ligands, as the most information-rich features, *i.e.*, the Watson–Crick base pairs, are buried deeply within the duplex core.

Pioneering probe technologies capable of targeting double-stranded DNA (dsDNA) include minor groove-binding pyrrole-imidazole polyamides,<sup>2,3</sup> and major groove-binding triplex-forming ONs (TFOs)<sup>4,5</sup> and peptide nucleic acids (PNAs),<sup>5,6</sup> which access nucleotide-specific features from the duplex grooves (Fig. 1). However, these approaches have a limited target scope; polyamides are generally directed towards short dsDNA regions as binding- and shape-complementarity is compromised when longer regions are targeted, whereas stable triplex formation by TFOs and PNAs is normally restricted to regions with extended polypurine tracts. While polyamides can be linked to target longer dsDNA regions,<sup>7</sup> and strategies reducing the polypurine requirement for triplex formation have been developed<sup>5,8–13</sup> – such as alternate strand TFOs,<sup>11,12</sup> the use of modified nucleotides capable of recognizing interruption sites,<sup>10,13</sup> or joining two PNA strands *via* a non-targeting linker (bis-PNA)<sup>8</sup> and extending the Watson–Crick binding end of the bis-PNA to yield tail-clamp (tc)-PNA<sup>9</sup> (Fig. 1) – this increases the complexity of the probes without fully addressing the challenges.

Recently, CRISPR–Cas9 has received considerable attention for DNA-targeting applications (Fig. 1). While the potential of CRISPR–Cas9 and related approaches is clear, current limitations must be acknowledged; CRISPR-based systems often suffer from immunogenicity, delivery challenges and specificity

concerns.<sup>14</sup> Efforts to address these limitations are needed for CRISPR to become a universal tool for *in vivo* applications.

Development of simpler, ON-based probe technologies capable of sequence-unrestricted recognition of double-stranded DNA therefore remains an important endeavor. Towards this end, numerous strand-invading probes have been explored, *i.e.*, probes capable of breaking existing Watson–Crick base pairs of dsDNA regions to form new, more stable base pairs between probe strands and complementary DNA (cDNA) regions. Notable examples of strand-invading probes (Fig. 2) – among others<sup>15–17</sup> – include:

- $\gamma$ PNA, in which the electrostatically neutral and achiral *N*-(2-aminoethyl)glycine backbone of PNA is additionally modified at the  $\gamma$ -position with a chirality-inducing moiety such as an (*R*)-methyl group (<sup>Me</sup> $\gamma$ PNA), resulting in preorganization of the strand and increased cDNA affinity compared to conventional PNA.<sup>18</sup> To increase the aqueous solubility and reduce aggregation of  $\gamma$ PNAs, (*R*)-hydroxyl (<sup>Se<sub>r</sub></sup> $\gamma$ PNAs) or (*R*)-diethylene glycol moieties (mini-PEG (MP), <sup>MP</sup> $\gamma$ PNAs) have been introduced at the  $\gamma$ -position.<sup>19,20</sup>

- acpcPNA, *i.e.*, single-stranded PNAs featuring an (2'*R*,4'*R*)-proline/(1*S*,2*S*)-2-aminocyclopentanecarboxylic (acpc) backbone. The pyrrolidine ring and  $\beta$ -amino acid of this  $\alpha/\beta$ -dipeptide backbone preorganize the strand for efficient dsDNA invasion, while the inherent bulkiness of the acpcPNA moieties prevent self-hybridization.<sup>21</sup>

- Locked nucleic acids (LNAs), *i.e.*, single-stranded ONs that are extensively modified with 2'-*O*,4'-*C*-methylene- $\beta$ -*D*-ribofuranosyl monomers. The dioxabicyclo[2.2.1]heptane skeleton locks the furanose ring in an *N*-type conformation (*C*3'-*endo*, pseudorotational phase angle  $P = 17^\circ$ ), leading to structural preorganization of the strand for target recognition.<sup>22</sup> While primarily renowned for their RNA-binding properties,<sup>23</sup> LNAs have been shown to recognize mixed-sequence dsDNA regions under conditions in which the targets are particularly accessible.<sup>24–27</sup>

- Pseudocomplementary DNA and PNA (pcDNA/pcPNA),<sup>28–31</sup> *i.e.*, short DNA or PNA duplexes that are modified with pseudocomplementary base pairs such as 2-thiouracil and 2,6-diaminopurine. Steric clashes between the pseudocomplementary base pairs perturb and weaken the hydrogen bonding between these moieties, resulting in destabilized probe duplexes, while the pseudocomplementary nucleobases form stable base pairs with

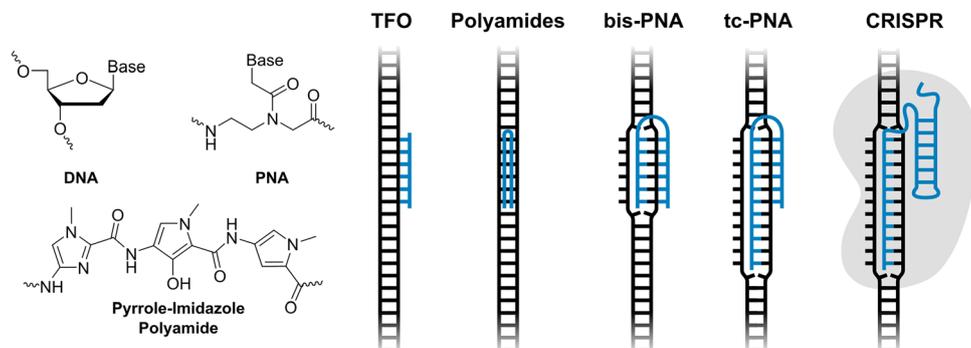


Fig. 1 Overview of pioneering DNA-targeting strategies.



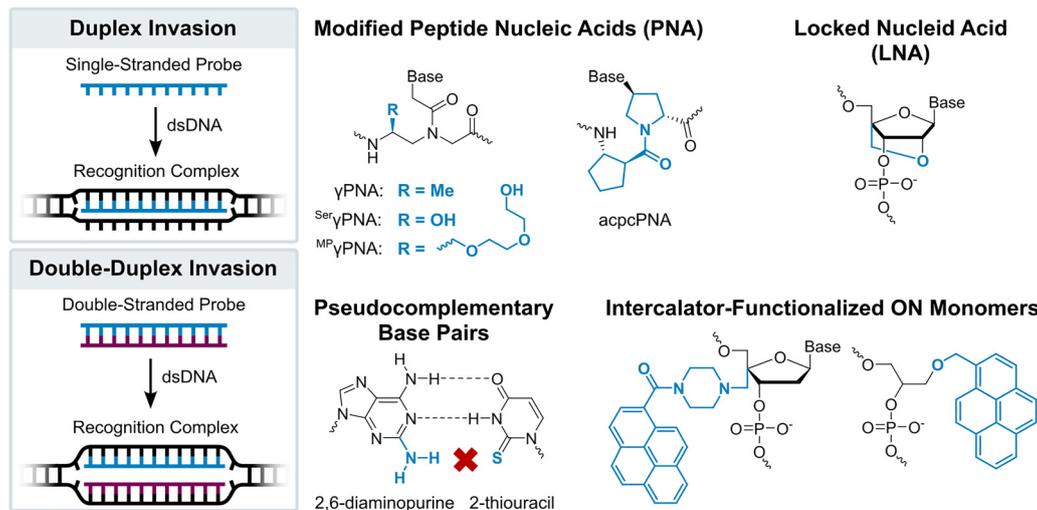


Fig. 2 Illustrations of DNA recognition mechanisms and overview of chemistries used in strand-invading probes.

canonical nucleobases. The stability difference between the double-stranded probes and the corresponding duplexes between individual probe strands and cDNA provides the driving force for double-duplex invasion of dsDNA target regions under certain conditions.

– Heteroduplexes between intercalator-modified ONs and complementary PNA, RNA, or LNA strands.<sup>32–38</sup> These approaches rely on intercalators typically being accommodated less well in PNA/DNA duplexes or A-type (RNA-like) duplexes,<sup>39,40</sup> than B-type (DNA-like) duplexes. As a result, the heteroduplexes are less stable than the corresponding duplexes between individual probes strands and cDNA. Akin to pcDNA and pcPNA, this stability difference provides the driving force for recognition of dsDNA targets.

A key advantage of strand-invading probes is that dsDNA regions of any sequence composition, in principle, can be targeted. However, single-stranded probes that are extensively modified with affinity-enhancing modifications tend to self-hybridize,<sup>35–38,41</sup> which can compromise target binding (acpcPNA being a notable exception).<sup>21</sup> Double-stranded probes, in turn, must be engineered to denature easily whilst maintaining high cDNA affinity.

For the past twenty years, our laboratory has pursued development of an alternative class of dsDNA-targeting probes relying on double-duplex invasion, which we have named Invader probes. These are double-stranded oligonucleotide probes featuring intercalator-functionalized nucleotides that are specifically arranged to promote destabilization of the probe duplex, while individual strands exhibit very high affinity to cDNA (Fig. 3). This account will describe the discovery, principles, and applications of Invader probes.

## Discovery of the Invader principle

Following its pioneering work on conformationally restricted and affinity-enhancing locked nucleic acid (LNA) and  $\alpha$ -L-LNA nucleotide monomers in the late 1990s and early 2000s (Fig. 4),<sup>42,43</sup>

Prof. Jesper Wengel's laboratory developed an interest in using functionalized oligonucleotides for nucleic acid-based Ångström-scale engineering.<sup>44</sup> This, in turn, fueled an interest in N2'-functionalized derivatives of 2'-amino-LNA<sup>45,46</sup> and 2'-amino- $\alpha$ -L-LNA monomers<sup>47,48</sup> (Fig. 4) as substitution of the oxygen atom for a nitrogen atom allows for introduction of additional, precisely positioned moieties, such as fluorescent pyrenes.<sup>49</sup> One of us (PJH) was charged with devising a synthesis of N2'-functionalized 2'-amino- $\alpha$ -L-LNA thymine monomers, after an initial attempt had unexpectedly yielded the corresponding 2'-amino-xylo-LNA monomers<sup>50</sup> instead.

### Synthesizing N2'-pyrene-functionalized 2'-amino- $\alpha$ -L-LNA monomers

Following extensive experimentation, a long but viable synthetic route to the phosphoramidite of 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA thymine monomer **L** (*i.e.*, the first Invader building block, *vide infra*) was developed using inexpensive diacetone- $\alpha$ -D-glucose as a starting material (~3% overall yield, 18 steps, Scheme 1).<sup>51,52</sup> Key steps included: (i) introduction of a C2-azido group *prior* to installation of the nucleobase, (ii) Vorbrüggen glycosylation affording an inseparable anomeric mixture of  $\alpha$ - and  $\beta$ -nucleosides (~55 : 45 ratio), (iii) separation of subsequent  $\alpha$ -L-*ribo*- and  $\beta$ -L-*ribo*-configured bicyclic nucleoside intermediates, (iv) selection of a suitable protecting group to prevent intramolecular Michael addition of the C2'-amino group to the C6-position of thymine and formation of tetracyclic 5,6-dihydrothymine derivatives, and (v) N2'-functionalization of the DMTr-protected key amino alcohol intermediate. Other intercalators were subsequently introduced in this manner.

Several attempts were made to shorten the route and avoid formation of anomeric nucleoside mixtures by converting nucleoside intermediates of the general type **1** into intermediate **2** (Scheme 2).<sup>51</sup> However, these attempts were unsuccessful, either resulting in undesirable formation of the corresponding O2,O2'-anhydronucleosides (eventually leading to the undesirable



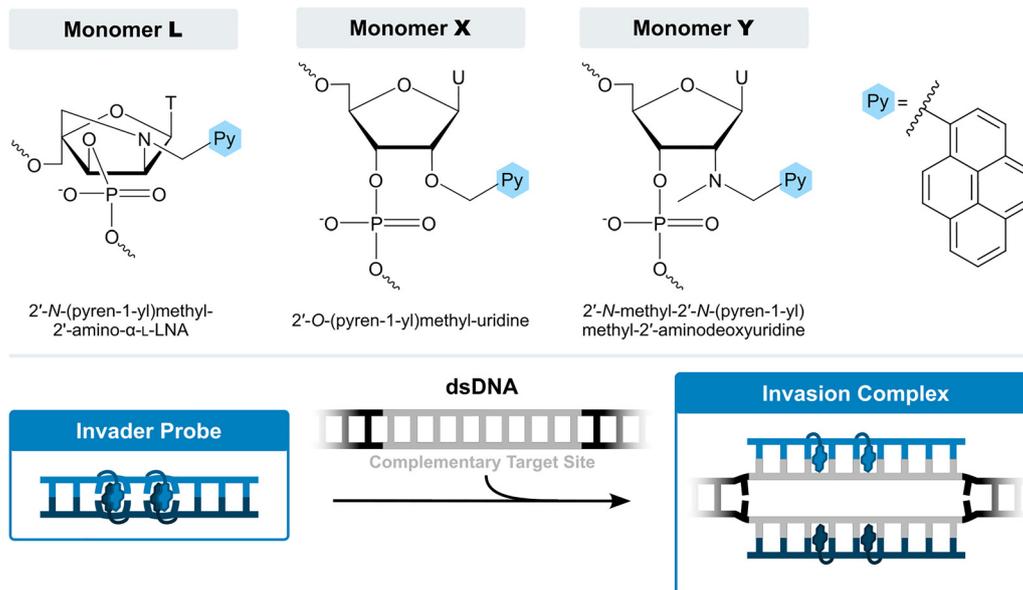


Fig. 3 Structures of key Invader monomers and illustration of Invader probe-mediated invasion of dsDNA.

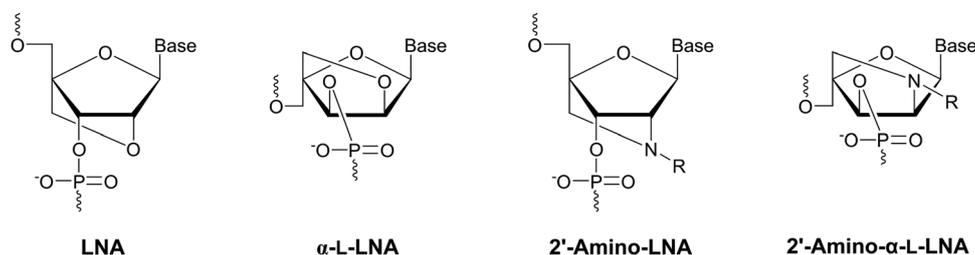
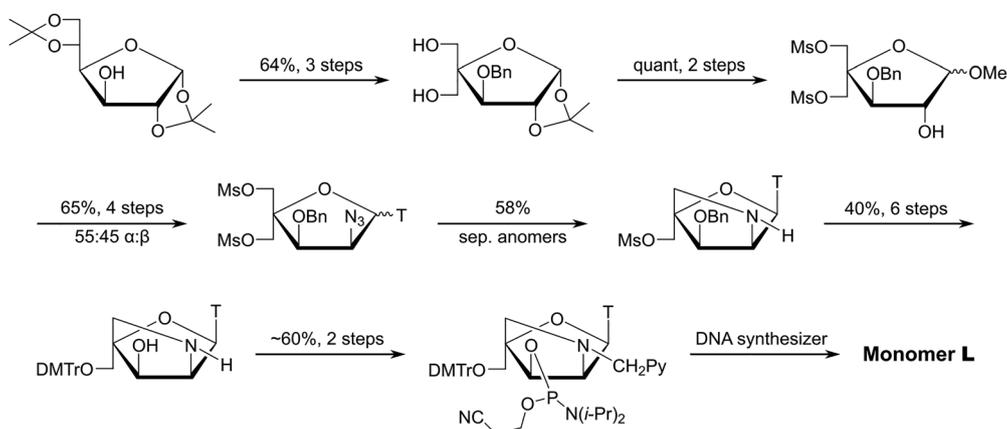


Fig. 4 Structures of LNA,  $\alpha$ -L-LNA, and N2'-functionalized 2'-amino-LNA and 2'-amino- $\alpha$ -L-LNA monomers.



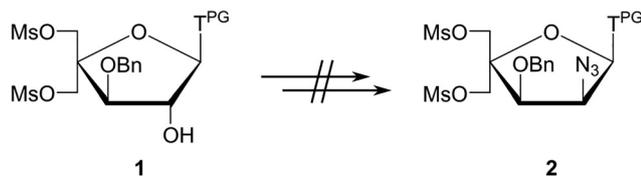
Scheme 1 Outline of the synthetic route to the phosphoramidite of 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA thymine monomer L.

downstream formation of 2'-amino-xylo-LNA nucleosides<sup>50</sup>) or no reaction (presumably due to the steric hindrance at the C2' ( $\alpha$ )-position stemming from the 1,3-*cis* orientation of the thymine moiety and the C3'-benzyloxy group).

In parallel, we developed a synthetic route to the corresponding N6-benzoyladenine analogue of monomer L, but this

proved even more challenging (<1% overall yield over 18 steps from diacetone- $\alpha$ -D-glucose, Scheme 3).<sup>53</sup> Although the key C2'-azido group could be successfully introduced with anchimeric assistance from the O2'-acetyl group after installation of the nucleobase – thereby avoiding formation of anomeric nucleoside mixtures – the route was marred by challenging





**Scheme 2** Unsuccessful attempts at shortening the route to key intermediate **2** in the synthesis of 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA thymine monomers.

downstream protecting group manipulations, which reduced the overall yield. Nonetheless, enough amidite was obtained for preliminary characterization studies.

**Key point:** Synthetic routes to the corresponding phosphoramidites of 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers have been developed. However, the routes are long and low-yielding.

### Preliminary characterization of ONs modified with 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers

In an initial study, we were pleased to discover that 9-mer ONs modified with a single **L** monomer form exceptionally stable duplexes with cDNA as evidenced by increases in thermal denaturation temperatures ( $T_m$ s) of up to +15.5 °C, relative to the corresponding unmodified DNA duplexes (*i.e.*,  $\Delta T_m$ ).<sup>54</sup> A subsequent study<sup>52</sup> revealed that the extent of duplex stabilization significantly exceeds that of  $\alpha$ -L-LNA-T or unfunctionalized 2'-amino- $\alpha$ -L-LNA-T monomers ( $\Delta T_m$ s = up to +8.0 °C and +2.5 °C, respectively), indicating that the pyrene moiety has a significant stabilizing role that acts in addition to any affinity-enhancing effects of the bicyclo[2.2.1]heptane scaffold. In stark contrast, incorporation of a single 2'-*N*-ethyl-2'-amino- $\alpha$ -L-LNA thymine monomer – *i.e.*, a *de facto* replacement of the pyrene moiety for a methyl group – results in dramatic destabilization of DNA duplexes ( $\Delta T_m$ s as low as –12.0 °C),<sup>52</sup> further underscoring the stabilizing role of the pyrene moiety. Duplex stabilization has been attributed to the formation of strong  $\pi$ - $\pi$  stacking interactions between the intercalating pyrene moiety and flanking base pairs, a binding hypothesis that is additionally supported by the following observations:<sup>52,54</sup>

– **L**-modified ONs display prominent DNA selectivity, *i.e.*, duplexes between **L**-modified ONs and complementary

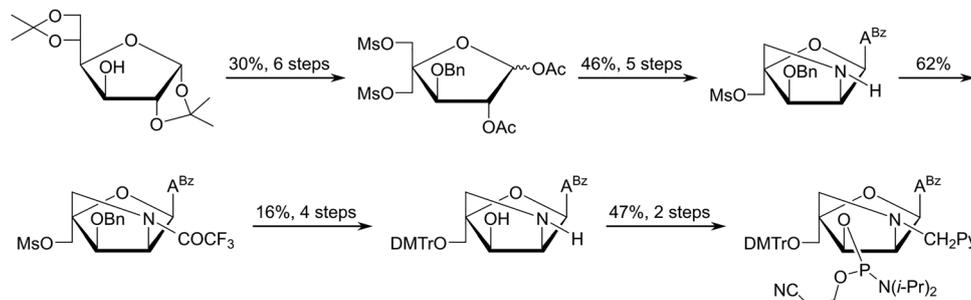
RNA (cRNA) are far less stabilized ( $\Delta T_m$  (ON:cDNA) –  $\Delta T_m$  (ON:cRNA) = 6.5–9.0 °C). Intercalators are known to favor B-type duplexes over A-type duplexes,<sup>40</sup> as the former are more elongated, with more efficient base pair overlaps for stronger stacking interactions with intercalators. Accordingly, DNA selectivity is often observed for ONs modified with intercalating pyrene moieties.<sup>32,55–58</sup>

– Compared to unmodified ONs, **L**-modified ONs display reduced discrimination of DNA strands with mismatched nucleotides opposite the pyrene-modified monomer. Thus, duplexes between a centrally **L**-modified 9-mer ON and DNA strands with mismatched C/G/T residues opposite of the **L** monomer displayed  $\Delta T_m$  values of –12.5/–5.5/–3.5 °C relative to the corresponding matched duplexes, respectively, as compared to  $\Delta T_m$  values of –16.5/–9.5/–17.0 °C for the corresponding unmodified ON.<sup>52</sup> This is likely because strong stacking interactions between the intercalating pyrene moiety and flanking mismatched base pairs compensate for the loss in stability from the distorted base pair.<sup>59,60</sup>

– Duplex formation between **L**-modified ONs and cDNA results in bathochromic and hypsochromic shifts of pyrene absorption maxima, which are indicative of strong electronic interactions between an intercalating pyrene moiety and flanking nucleobases.<sup>61</sup>

Closer inspection of the structure of monomer **L** reveals that the attachment points of the nucleobase and pyrene moieties are restricted relative to each other due to the rigid nature of the 2-oxo-5-azabicyclo[2.2.1]heptane scaffold and the short methylene linker used to connect the pyrene. This promotes  $\pi$ - $\pi$  stacking between the nucleobase and pyrene moiety, which accordingly is preorganized to intercalate. Molecular modelling studies provided further support for this hypothesis (Fig. 5). The lowest energy structure of a singly **L**-modified DNA duplex featured an intercalating pyrene moiety engaging in extensive  $\pi$ - $\pi$  stacking with the thymine of monomer **L** and the 3'-flanking nucleobase and, to a lesser extent, the nucleobases on the complementary strand across of these moieties.<sup>52</sup> Consistent with this, duplex stabilization is more pronounced when the **L** monomer is flanked by 3'-purines with larger  $\pi$ -stacking surfaces.<sup>52,54</sup>

**Key point:** ONs modified with 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers exhibit exceptional cDNA affinity due to 3'-directed intercalation of the pyrene moiety.



**Scheme 3** Outline of the synthetic route to the phosphoramidite of the adenine analogue of monomer **L**.



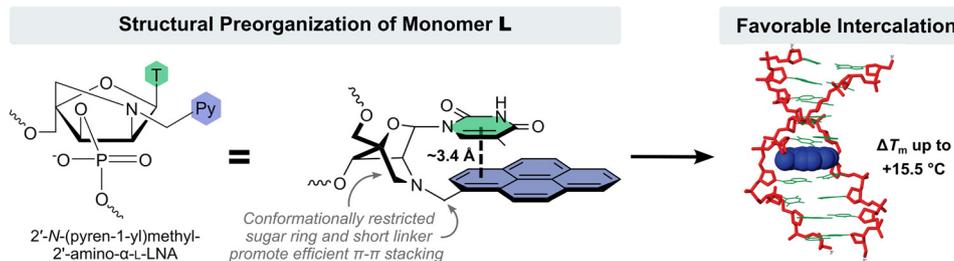


Fig. 5 Structural preorganization of monomer L resulting in forced intercalation of the pyrene moiety. Image adapted from ref. 52 with permission from the American Chemical Society, copyright 2009.

### DNA duplexes modified with different interstrand zippers of 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers

After establishing the high cDNA-affinity of L-modified ONs, a series of 9-mer duplexes with different “interstrand zipper arrangements” of L monomers were evaluated (Fig. 6),<sup>54</sup> to explore novel structural motifs potentially capable of interstrand communication for nucleic acid-based Ångström-scale engineering.<sup>44</sup>

Duplexes with  $-3$ ,  $-1$ , and  $+4$  interstrand zipper arrangements of L monomers were found to be strongly stabilized, with  $\Delta T_m$  values between  $+15.5$  °C and  $+25.0$  °C (Fig. 7a). To assess if two L monomers contribute independently to duplex stability, the term deviation from additivity (DA) was introduced and calculated as follows:  $DA_{\text{ONX:ONY}} = \Delta T_m(\text{ONX:ONY}) - \Delta T_m(\text{ONX:cDNA}) - \Delta T_m(\text{ONY:cDNA})$ , where ONX:ONY represents a duplex with an interstrand zipper arrangement of monomers.<sup>62</sup> For duplexes with  $-3$ ,  $-1$ , and  $+4$  zippers of L monomers, the DA values ranged from  $-8.0$  °C to  $+2.5$  °C, indicating that the L monomers largely impact DNA duplex stability independently from each other with contributions ranging from moderately less-than-additive to slightly more-than-additive (Fig. 7a). Crucially, a DNA duplex with a  $+1$  zipper of L monomers was found to be considerably less stable with a  $\Delta T_m$  value of  $+2.5$  °C and a DA value of  $-27.0$  °C (Fig. 7a). The corresponding duplex with a  $+2$  zipper of L monomers was also labile but exhibited less prominent destabilizing interference between the two L monomers ( $\Delta T_m \sim 0$  °C,  $DA = -13.5$  °C; Fig. 7a). Subsequent characterizations of DNA duplexes with a broader range of interstrand zipper arrangements of L monomers revealed that the prominent destabilizing interference is unique to  $+1$  zippers.<sup>63</sup> As will be discussed in subsequent

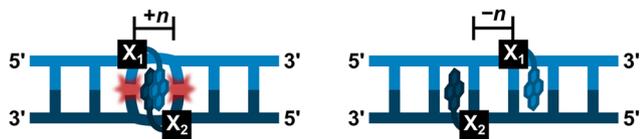


Fig. 6 Illustration of interstrand zipper arrangements. The number  $n$  describes the distance measured in number of base pairs and has a positive value if a monomer is shifted toward the 5'-side of its own strand relative to a second reference monomer on the other strand or a negative value if a monomer is shifted toward the 3'-side of its own strand relative to a second reference monomer on the other strand.

sections, this is the key structural motif that forms the basis for dsDNA-targeting Invader probes.

DNA duplexes with  $+1$  (or any other) interstrand zippers of  $\alpha$ -L-LNA-T or unfunctionalized 2'-amino- $\alpha$ -L-LNA-T monomers do not display comparable destabilizing interference (shown later in Table 1), reinforcing the notion that the distinctive nature of DNA duplexes with  $+1$  zippers of L monomers, primarily is due to the two pyrene moieties, rather than the proximity of two bicyclo[2.2.1]heptane scaffolds.<sup>62</sup>

The distinctive nature of DNA duplexes with  $+1$  zippers of L monomers was further underscored by the presence of prominent pyrene excimer signals, centered at 495 nm, in the steady-state fluorescence emission spectra of these duplexes (Fig. 7b).<sup>54,63</sup> These signals imply a coplanar arrangement of pyrene moieties with an interplanar separation of  $\sim 3.4$  Å.<sup>64</sup> In contrast, duplexes with other interstrand zipper arrangements exhibited negligible, if any, emission in this region (Fig. 7b).

Preliminary force field simulations of a 9-mer DNA duplex with a  $+1$  zipper of L monomers indicated that the spatial separation between the two pyrene moieties is too large to account for excimer formation across the major groove (estimated shortest distance  $\sim 12$  Å).<sup>54</sup> Subsequent, more extensive molecular modeling studies suggested that simultaneous 3'-directed intercalation of both pyrene moieties is possible (Fig. 7c).<sup>62</sup> The resulting stacking arrangement is consistent with excimer formation and perturbs neighboring base pairs and partially elongates and unwinds the probe duplex, consistent with reduced duplex stability. It should be noted that the lowest energy structure from this simulation only featured one intercalating pyrene, while the other pyrene projected into the major groove.<sup>62</sup> While this structure also could explain the low duplex stability (*i.e.*, placement of a non-polar pyrene perturbing the hydration spine in the major groove), it does not account for the observed pyrene excimer emission.

These conclusions were additionally substantiated by NMR studies on a 13-mer DNA duplex with a central  $+1$  zipper of L monomers. The  $^1\text{H}$  NMR signals from nucleotides proximal to the L monomers exhibited progressive broadening, while signals of protons close to the pyrene moieties were entirely absent. This indicates a significant level of perturbation and flexibility near the  $+1$  zipper.<sup>62</sup> Similar structural consequences were previously reported for NMR structures of DNA duplexes with adjacent incorporations of pyrene-modified non-nucleotide monomers.<sup>65</sup>



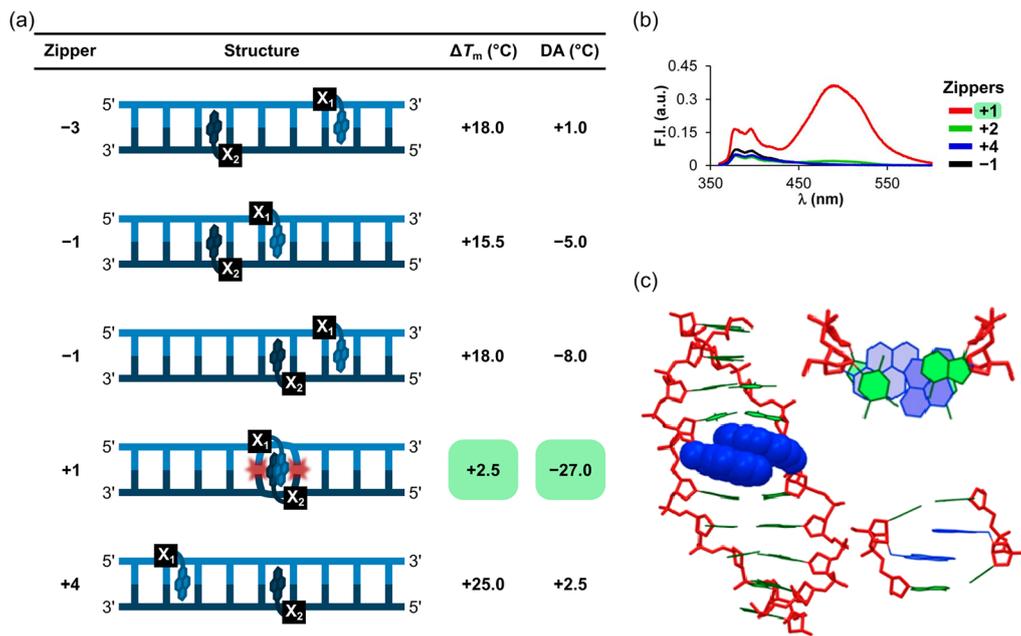


Fig. 7 Characteristics of DNA duplexes modified with different interstrand zipper arrangements of L monomers. (a)  $\Delta T_m$  and DA values. (b) Steady-state fluorescence emission spectra ( $\lambda_{\text{ex}} = 350$  nm) of DNA duplexes with +1, +2, +4, or -1 zippers of L monomers. (c) Energy-minimized structure of a DNA duplex with a +1 zipper of L monomers. Figure adapted from ref. 62 with permission from the American Chemical Society, copyright 2013.

The perturbations in the vicinity of the +1 zipper of L monomers are consistent with a violation of the nearest-neighbor exclusion principle (NNEP).<sup>66,67</sup> The principle posits that intercalator densities exceeding one intercalator per two base pairs are unfavorable in DNA duplexes. This is due to constraints in local helix expandability (each intercalation event expands the duplex by  $\sim 3.4$  Å) and because stabilizing stacking interactions between flanking base pairs and a first intercalating moiety are perturbed upon intercalation of a second moiety.<sup>68,69</sup>

Duplexes with +1 zippers of L monomers exhibit a local density of two intercalators per two base pairs – which represents a major violation of the NNEP – resulting in a labile duplex (e.g., see +1 zipper structure in Fig. 7a). Duplexes with +2 zippers of L monomers or 0 zippers between the L monomer and its corresponding adenine derivative have a local density of two intercalators per three base pairs according to our binding model (e.g., see the +2 zipper structure in Fig. 7a), which represents a more moderate violation of the NNEP, consistent with the less pronounced destabilizing interference.<sup>54,62</sup> Duplexes with other interstrand zipper arrangements of L monomers are very stable since the NNEP is not violated (local intercalator density is – at most – one intercalator per two base pairs).<sup>54,63</sup> Duplexes between ONs modified with interspersed L monomers and cDNA are also very stable since the local intercalator density never exceeds the NNEP limit.

**Key point:** DNA duplexes with a +1 zipper of L monomers are labile, presumably due to 3'-directed intercalation of both pyrene moieties into a shared region between neighboring base pairs, leading to unique destabilizing interference and prominent violation of the NNEP. Less prominent or no destabilizing interference between the pyrene moieties are observed in DNA

duplexes with other interstrand zipper arrangements of L monomers and in duplexes between ONs with non-sequential incorporations of L monomers and cDNA.

### Defining key terminology

For the remainder of this account, the structural motif described as a +1 interstrand zipper of pyrene-functionalized nucleotide monomers (and more generally, a +1 interstrand zipper of intercalator-functionalized nucleotide monomers) is referred to as an *energetic hotspot*. An oligonucleotide-based duplex featuring one or more energetic hotspots, is referred to as an *Invader probe*. An Invader probe featuring one or more energetic hotspots of L monomers, is specifically referred to as an *Invader LNA probe*.

### Initial dsDNA-recognition experiments using double-stranded probes with +1 zippers of 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers

The difference in stability between DNA duplexes with a +1 interstrand zipper of L monomers (i.e., Invader LNAs, more labile) versus duplexes between L-modified ONs and cDNA (i.e., probe-target duplexes, more stable) led us to stipulate that recognition of complementary dsDNA targets using Invader LNAs would be thermodynamically feasible and proceed *via* a double-duplex invasion process, provided that recognition kinetics are favorable.

In a preliminary proof-of-concept experiment, a pre-annealed 9-mer Invader LNA featuring a single +1 interstrand zipper of L monomers was incubated with an equimolar quantity or twofold excess of a pre-annealed, unmodified DNA duplex (mixed-sequence, complementary to the probe; Fig. 8).<sup>54</sup> A rapid



(<1 min) disappearance of the pyrene excimer signal was observed, indicating that the Invader LNA probe had dissociated to form two corresponding probe-target duplexes (Fig. 8). Importantly, the experiments were performed in  $T_m$  buffer (formulated to mimic physiologic salinity:  $[\text{Na}^+] = 110 \text{ mM}$ ) at temperatures significantly below the  $T_m$ s of the Invader LNA probe, the dsDNA target, or any of the probe-target duplexes that would form upon successful recognition, indicating that the recognition process might proceed *via* a sequential displacement (strand invasion) process rather than a simple dissociative (strand exchange) process.<sup>70</sup> Conversely, incubation of the Invader LNA probe with an equimolar quantity of a 9-mer DNA duplex differing in sequence at one or two positions relative to the probe (referred to as singly- or doubly-mismatched dsDNA targets despite being fully base-paired), resulted in much smaller decreases in excimer emission (Fig. 8).

Following our initial proof-of-concept study with Invader LNA probes,<sup>54</sup> dsDNA-recognition using a closely related approach based on so-called intercalating nucleic acids (INAs) was reported.<sup>71</sup> INAs are double-stranded ON probes with opposing bulges of non-nucleotidic 1-*O*-(1-pyrenylmethyl)-glycerol monomers (see Fig. 2, bottom right corner). Akin to Invader LNA probes, this design results in a partially unwound and destabilized probe duplex<sup>65</sup> whereas individual probe strands display high cDNA affinity, providing the driving force for dsDNA-recognition.

**Key point:** Invader LNA probes enable specific recognition of short, complementary mixed-sequence dsDNA at physiologically relevant salt concentrations.

### Invader LNAs with one or two energetic hotspots – preliminary insights into probe optimization

To further assess the viability of this approach for recognition of linear, mixed-sequence dsDNA targets, seven 13-mer Invader LNAs featuring one or two +1 zippers of L monomers were synthesized.<sup>63</sup> Importantly, similar results were observed, validating the initial findings. Thus, duplexes between individual

probe strands and cDNA were found to be significantly stabilized ( $\Delta T_m$  per modification ( $\Delta T_m/\text{mod}$ ) values between +7.0 °C and +11.8 °C). ONs with two L monomers as next-nearest neighbors displayed slightly less-than-additive increases in  $\Delta T_m$  values, presumably as the local intercalator density approached the NNEP limit (*i.e.*, two intercalators per four base pairs). Conversely, more-than-additive stabilization was observed when two L monomers were separated further apart. The duplex stabilization was found to be due to less unfavorable changes in entropy compared to formation of the corresponding unmodified DNA duplex, likely reflecting the conformationally restricted nature of the bicyclic 2'-amino- $\alpha$ -L-LNA scaffold, which forces the attached pyrene moiety into the duplex core. Duplexes between individual probe strands and cRNA were found to be far less stabilized ( $\Delta T_m/\text{mod}$  between +0.5 °C and +6.3 °C), consistent with the preliminary study.

Invader LNAs featuring a single energetic hotspot were found to be quite labile ( $\Delta T_m = -1.5$  °C to +3.5 °C), as in the preliminary study. The lability is due to a less favorable change in enthalpy, consistent with localized intercalator-mediated perturbation of base-pairing and violation of the NNEP. Invader LNAs with two hotspots were found to be quite stable, although  $T_m$  and  $\Delta G^{293}$  values (*i.e.*, the change in Gibbs free energy at 293 K upon duplex formation) provided conflicting indications ( $\Delta T_m = +10.5$  °C to +15.5 °C and  $\Delta \Delta G^{293} = 1-2 \text{ kJ mol}^{-1}$ , relative to the corresponding unmodified dsDNA).<sup>63</sup>

The free energy for recognition of isosequential dsDNA targets at 293 K by the different Invader LNA probes was estimated as follows:  $\Delta G_{\text{rec}}^{293} = \Delta G^{293} (5'\text{-ON:cDNA}) + \Delta G^{293} (3'\text{-ON:cDNA}) - \Delta G^{293} (5'\text{-ON:3'-ON}) - \Delta G^{293} (\text{dsDNA})$ , where 5'-ON:3'-ON denotes an Invader LNA. Invader LNAs featuring a single hotspot displayed  $\Delta G_{\text{rec}}^{293}$  values between  $-26 \text{ kJ mol}^{-1}$  and  $-19 \text{ kJ mol}^{-1}$ , while probes with two hotspots were more strongly activated for dsDNA-recognition ( $\Delta G_{\text{rec}}^{293}$  values between  $-50 \text{ kJ mol}^{-1}$  and  $-34 \text{ kJ mol}^{-1}$ ). Probes with two non-consecutive hotspots were found to display more favorable  $\Delta G_{\text{rec}}^{293}$  values than probes with two sequential hotspots. Similar conclusions were drawn from the  $T_m$ -based equivalent of the  $\Delta G_{\text{rec}}$  term, *i.e.*, thermal advantage (TA) calculated as  $\text{TA} = T_m$

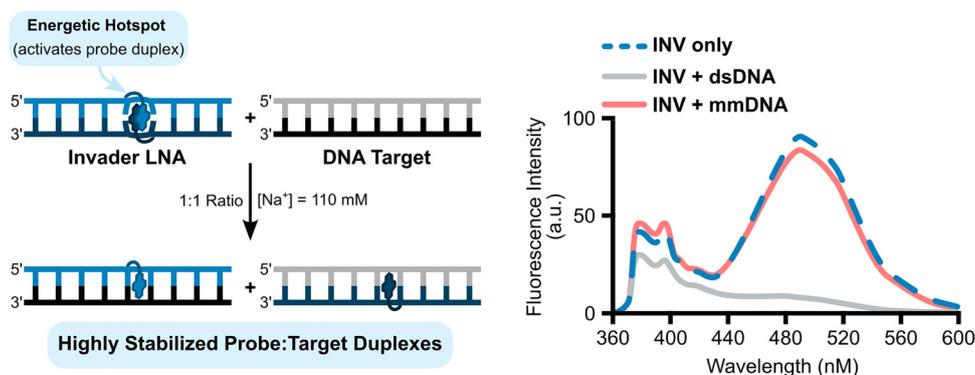


Fig. 8 Illustrations of Invader LNA-mediated recognition of dsDNA and steady-state fluorescence emission spectra of an Invader LNA (INV) probe incubated with a complementary dsDNA (INV + dsDNA) or doubly-mismatched DNA (INV + mmDNA). Figure adapted from ref. 54 with permission from the Royal Society of Chemistry.



$(5'-\text{ON}:\text{cDNA}) + T_m(3'-\text{ON}:\text{cDNA}) - T_m(5'-\text{ON}:3'-\text{ON}) - T_m(\text{dsDNA})$  (note  $\text{TA} = -\text{DA}$ ).

The impact of buffer ionic strength on the driving force for dsDNA-recognition was also assessed; Invader LNAs with two hotspots exhibited increasingly favorable  $\Delta G_{\text{rec}}^{293}$  values with decreasing ionic strength, whereas  $\Delta G_{\text{rec}}^{293}$  values for single hotspot Invader LNAs remained essentially unaffected by changes in ionic strength. These trends arise because individual probe strands display additionally increased cDNA affinity relative to unmodified strands as ionic strength is decreased, whereas the relative stability of Invader probes only is minimally impacted. Consequently, recognition of dsDNA targets with densely modified Invader LNAs is expected to become more thermodynamically favorable in environments with low ionic strength, with potential important ramifications for recognition of chromosomal DNA.

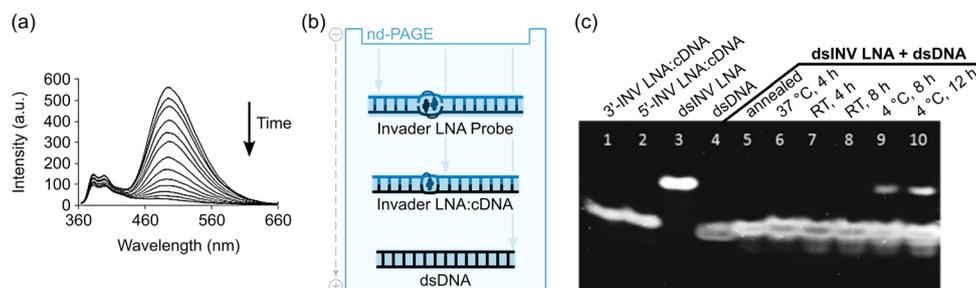
A subset of these 13-mer Invader LNAs were pre-annealed and then incubated with an equimolar quantity of the corresponding pre-annealed dsDNA targets at room temperature, *i.e.*, at an experimental temperature well below the  $T_m$ s of the dsDNA target (37.5 °C), Invader LNAs (36–53 °C), and probe-target duplexes (44.5–61.0 °C) that would form upon successful recognition. The recognition kinetics were monitored by observing the decrease in pyrene excimer emission (Fig. 9a). The recognition process was 50% complete ( $t_{50\%}$ ) within 29–48 min when experiments were performed in  $T_m$  buffer ( $[\text{Na}^+] = 110 \text{ mM}$ ). Only minor differences were observed between the Invader LNAs despite their different  $\Delta G_{\text{rec}}^{293}$  values, suggesting that target recognition – in this instance – is not thermodynamically limited. Conversely, when a representative 13-mer single hotspot Invader LNA was incubated with dsDNA targets differing in sequence at one position, a slower and less pronounced decrease in the excimer signal was observed. This was consistent with the unfavorable thermodynamics for recognition of mismatched targets ( $\Delta G_{\text{rec}}^{293}$  values between  $-1 \text{ kJ mol}^{-1}$  and  $+8 \text{ kJ mol}^{-1}$ ) and suggested that Invader LNA-mediated recognition of mixed-sequence dsDNA targets is highly specific.

Invader LNA probes featuring two sequential hotspots were found to result in significantly faster dsDNA-recognition than single hotspot Invader LNAs in high salt ( $[\text{Na}^+] = 710 \text{ mM}$ ) or

KMST buffers (pH 7.2, 140 mM KCl, 10 mM  $\text{MgCl}_2$ , 1 mM spermine and 40 mM Tris-Cl) ( $t_{50\%} = 126/31 \text{ min}$  and  $158/19 \text{ min}$ , for singly/doubly modified Invader LNAs in high salt and KMST buffers, respectively). These findings not only suggested that dsDNA-recognition is possible across a broad range of conditions but also hinted at preferred probe designs.

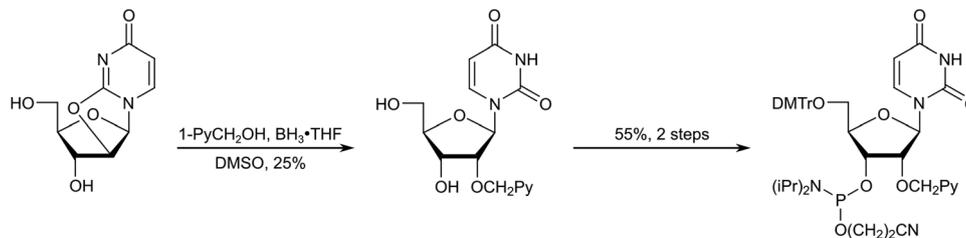
To validate the results from the fluorescence-based assay, electrophoretic mobility shift assays (EMSAs) entailing non-denaturing polyacrylamide gel electrophoresis (nd-PAGE) were conducted. The EMSAs rely on differences in the electrophoretic mobilities of Invader LNAs (lowest mobility; most elongated duplex due to double intercalation), isosequential dsDNA targets (highest mobility; most compact duplex), and probe-target duplexes (intermediate mobility and elongation) (Fig. 9b). A representative 13-mer single hotspot Invader LNA was incubated with an equimolar amount of the corresponding isosequential dsDNA target (4 h, at either 37 °C or room temperature, in a pH 7.2 HEPES buffer (70 mM HEPES, 15 mM  $\text{MgCl}_2$ , 15% sucrose, 0.15% spermine, 300 mM NaCl)), leading to the formation of probe-target duplexes as evidenced by a single band with the same mobility as the pre-annealed probe-target reference duplexes (compare lanes 6 and 7 with lanes 1, 2, and 5, Fig. 9c). No band corresponding to the Invader LNA (lane 3, Fig. 9c) was present, indicating these conditions facilitated complete recognition of the dsDNA target. Recognition proceeded, but remained incomplete, following incubation at 4 °C for 8–12 hours, as indicated by the presence of a faint band corresponding to the Invader LNA (compare lanes 9 and 10 with lane 3, Fig. 9c).

**Key point:** 13-Mer Invader LNA probes featuring hotspots of L monomers allow for fast and specific recognition of complementary dsDNA targets across different buffers – including buffers mimicking physiological conditions and high salt buffers. The difference in stability of the probe-target duplexes relative to the Invader LNAs and dsDNA targets renders recognition thermodynamically favorable. Invader LNAs with two hotspots display more favorable recognition thermodynamics and faster dsDNA-recognition kinetics than single hotspot Invader LNAs. The recognition process can be followed by monitoring the disappearance of the probe's pyrene excimer signal or by comparing relative mobilities in EMSAs.



**Fig. 9** dsDNA-recognition properties of Invader LNA probes. (a) Time course of steady-state fluorescence spectra upon addition of a 13-mer pre-annealed Invader LNA to pre-annealed isosequential dsDNA target. (b) Depiction of the different mobilities of Invader LNA probes, probe-target duplexes between Invader LNA strands and cDNA, and dsDNA in non-denaturing polyacrylamide gels. (c) Representative electrophoretogram demonstrating dsDNA-recognition by Invader LNA probes. Figure reproduced from ref. 63 with permission from the Royal Society of Chemistry.





Scheme 4 Outline of synthetic route to phosphoramidite of monomer X.

## Second-generation Invader probes

### Identifying and synthesizing more readily available Invader LNA mimics

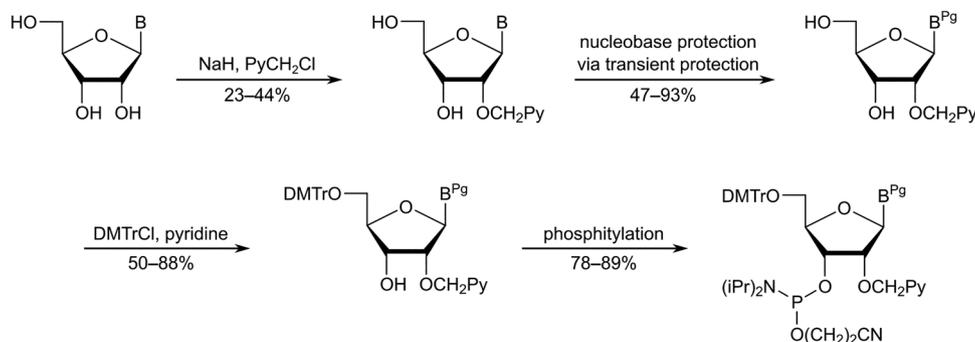
The initial studies highlighted the potential of Invader LNAs for sequence-unrestricted recognition of dsDNA targets. However, a comprehensive evaluation of these probes was not feasible due to the complex synthesis and limited supply of the necessary N2'-pyrene-functionalized 2'-amino- $\alpha$ -L-LNA phosphoramidites (Schemes 1 and 3). Consequently, our attention shifted to identifying and obtaining more readily available structural and functional mimics of N2'-pyrene-functionalized 2'-amino- $\alpha$ -L-LNA monomers. A survey of the existing literature at the time (approximately 2005–2010) suggested O2'-pyrene-functionalized RNA and N2'-pyrene-functionalized 2'-N-methyl-2'-amino-DNA monomers as potential candidates. ONs modified with 2'-O-(pyren-1-yl)methyluridine monomer **X**<sup>57</sup> or 2'-N-methyl-2'-N-(pyren-1-yl)methyl-2'-aminodeoxyuridine monomer **Y**<sup>58</sup> (Fig. 3) had been shown to exhibit high cDNA affinity. Moreover, the pyrene moiety of monomer **X** had been shown to intercalate in a 3'-directed manner upon hybridization with cDNA,<sup>72</sup> in a similar manner as proposed for monomer **L** (Fig. 5). Although unknown to us at the time, the high cDNA affinity bestowed by the pyrene moieties of monomers **L**, **X**, and **Y**, is likely linked to their sugar rings adopting *South*-type conformations. The next task then involved developing convenient synthetic routes to the corresponding phosphoramidites.

Yamana and co-workers reported a synthetic route to the corresponding phosphoramidite of monomer **X**.<sup>73</sup> The approach entailed (i) O3',O5'-ditritylation of uridine, (ii) installation of the

pyrene moiety at the O2'-position *via* Williamson's ether synthesis, (iii) detritylation, (iv) O5'-dimethoxytritylation, and (v) O3'-phosphitylation. In our hands, the O3',O5'-ditritylation only proceeded with modest regioselectivity and yield, and the protecting group manipulations following the O2'-functionalization were cumbersome. We therefore set out to develop an alternative route (Scheme 4).<sup>74</sup> Specifically, trialkyl borate-mediated opening<sup>75</sup> of O2,O2'-anhydrouridine using 1-pyrenemethanol in anhydrous DMSO afforded the O2'-pyrene-functionalized key intermediate in modest but consistent yields despite challenging column purifications. Subsequent O5'-dimethoxytritylation and O3'-phosphitylation, as reported by Yamana, afforded the target phosphoramidite in an overall yield of  $\sim$ 12% over five steps from uridine, a modest improvement over the original route. Other intercalators were subsequently installed in an equivalent manner.<sup>76–78</sup>

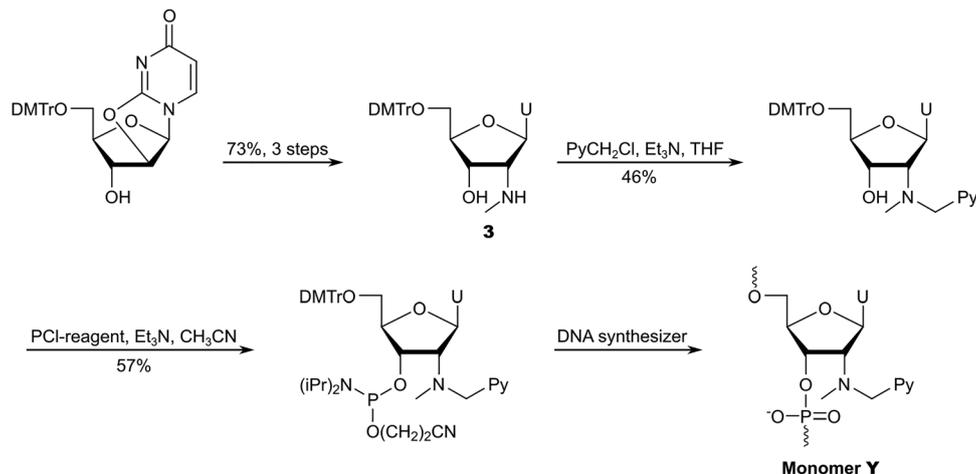
The protected adenine, cytosine, and guanine analogues were synthesized in a different manner as reported by Yamana and us (Scheme 5).<sup>79,80</sup> The free nucleosides were first O2'-pyrenylmethylated in modest regioselectivity and yield using Williamson's approach, followed by *N*-acylation of the exocyclic amines *via* transient protection protocols, O5'-dimethoxytritylation, and O3'-phosphitylation. This reaction sequence afforded the necessary phosphoramidites in overall yields ranging from 9% to 12% over four steps from the free nucleosides.

The Wengel group reported the synthesis of the corresponding phosphoramidite of monomer **Y** from 5'-O-dimethoxytrityl-2'-aminouridine in approximately 7% overall yield over five steps.<sup>58</sup> The route entailed N2'-pyrene-functionalization *via* reductive

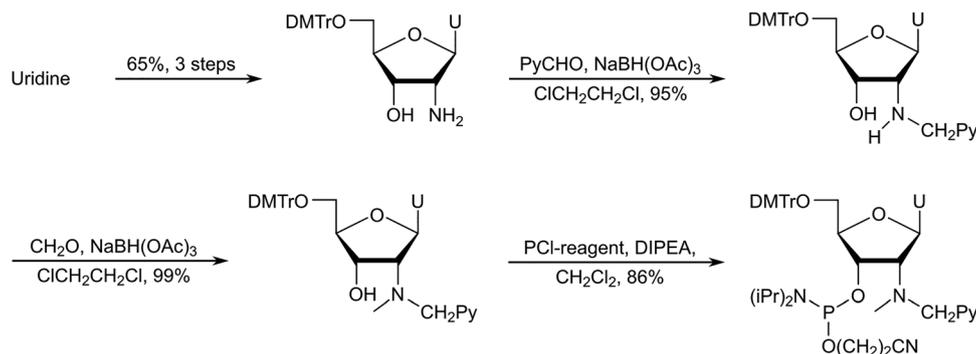


Scheme 5 Outline of synthetic route to the phosphoramidites of the protected adenine, cytosine and guanine analogues of monomer **X**. Pg = protecting group.





Scheme 6 Outline of our initial synthetic route to the phosphoramidite of monomer Y.



Scheme 7 Improved synthetic route to the phosphoramidite of monomer Y.

amination, O5'-detritylation, N2'-methylation under acidic conditions, O5'-dimethoxytritylation, and O3'-phosphitylation.

To reduce the number of protection/deprotection steps, we decided to use nucleoside **3**<sup>74</sup> as a starting material instead, which is obtained from 5'-O-dimethoxytrityl-2',2'-anhydrouridine in ~73% yield over three steps (Scheme 6).<sup>81</sup> Direct N2'-alkylation of **3** using 1-(chloromethyl)pyrene afforded the desired product in moderate yield, whereas reductive amination proved unsuccessful due to the formation of cyclic N2',O3'-hemiaminal ethers.<sup>74</sup> Subsequent O3'-phosphitylation provided the target phosphoramidite in 10% overall yield from uridine over seven steps.

After establishing monomer **Y** as a valuable building block for second-generation Invader probes (*vide infra*), we developed an improved route to the corresponding phosphoramidite (Scheme 7).<sup>82</sup> Using 5'-O-dimethoxytrityl-2'-amino-2'-deoxyuridine as a starting material<sup>83</sup> – which was obtained from uridine in three steps – we performed consecutive reductive aminations, first introducing the large N2'-substituent, then the N2'-methyl group. Improved O3'-phosphitylation afforded the phosphoramidite of monomer **Y** in ~52% yield from uridine over six steps,<sup>82</sup> representing a significant improvement over the initial route. This route also allowed for the introduction of other intercalators at the N2'-position.<sup>82</sup>

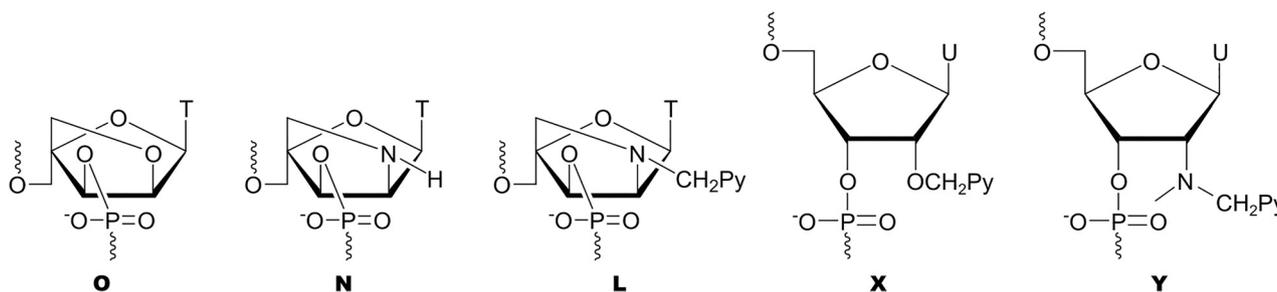
**Key point:** 2'-O-(Pyren-1-yl)methylribonucleoside and 2'-amino-2'-deoxy-2'-N-methyl-2'-N-(pyren-1-yl)methyluridine phosphoramidites can be obtained in far fewer steps and significantly higher overall yield than N2'-pyrene-functionalized 2'-amino- $\alpha$ -L-LNA phosphoramidites.

### Characterizing ONs modified with X or Y monomers

The impact of the **X** and **Y** monomers on duplex stability was examined alongside the  $\alpha$ -L-LNA-T (monomer **O**), unfunctionalized 2'-amino- $\alpha$ -L-LNA-T (monomer **N**), and 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA-T (monomer **L**) monomers.<sup>62,74</sup> Importantly, duplexes between **X**- or **Y**-modified ONs and cDNA were found to be only slightly less stable than the corresponding **L**-modified duplexes, but more stable than the corresponding **O**- and **N**-modified duplexes (Table 1). The stability trends of DNA duplexes with different interstrand zipper arrangements of monomers **X** and **Y** were found to closely mirror those of corresponding **L**-modified duplexes. Thus, DNA duplexes with -3, -1, and +4 interstrand zippers of **X** or **Y** monomers exhibit  $T_m$  values consistent with near-additive contributions from the two monomers. Duplexes with +2 zippers display  $T_m$ s indicative of moderately destabilizing interference between the two monomers, while duplexes with +1 zippers are as labile and with as strongly



**Table 1**  $\Delta T_m$  and TA values for Invader probe duplexes featuring O, N, L, X, or Y monomers in +4 or +1 zippers, and  $\Delta T_m$  values for duplexes between probe strands and cDNA



Zipper	Duplex	$\Delta T_m$ (°C)					
		<u>B</u> =	O	N	L	X	Y
+4	5'-GBG ATA TGC 3'-CAC TAB ACG		+11.0	+1.0	+25.5	+16.5	+19.5
	5'-strand vs. cDNA		+2.5	-2.0	+7.0	+5.0	+5.0
	3'-strand vs. cDNA		+8.0	+2.5	+15.5	+11.5	+13.0
	TA (°C) =		-0.5	-0.5	-2.5	0.0	-1.5
+1	5'-GTG ABA TGC 3'-CAC TAB ACG		+16.0	-5.5	+2.5	-2.0	-2.0
	5'-strand vs. cDNA		+6.0	+0.5	+14.0	+12.5	+14.0
	3'-strand vs. cDNA		+8.0	+2.5	+15.5	+11.5	+13.0
	TA (°C) =		-2.0	+8.5	+27.0	+26.0	+29.0

destabilizing interference as the corresponding Invader LNAs (compare  $T_m$  and TA values, Table 1).

The similar thermal denaturation characteristics of the X-, Y-, and L-modified duplexes indicated to us that the pyrene moieties have similar binding modes, *i.e.*, stabilizing intercalation in duplexes between the pyrene-modified ONs and cDNA, and destabilizing interference and violation of the NNEP in DNA duplexes with +1 zippers. The following observations corroborated this further:

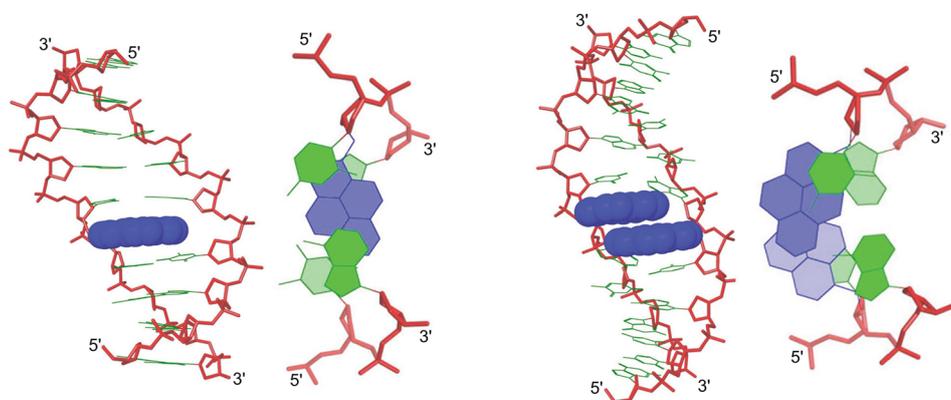
(i) X- and Y-modified ONs display similar selectivity for cDNA over cRNA as L-modified ONs<sup>74</sup>

(ii) The increased stability of duplexes between X-modified ONs and cDNA is also due to less unfavorable changes in entropy upon duplex formation<sup>76</sup>

(iii) X- and Y-modified ONs also display reduced discrimination of DNA strands with mismatched nucleotides positioned opposite the pyrene-functionalized monomer<sup>74</sup>

(iv) Hybridization of X- and Y-modified ONs with cDNA also results in bathochromic shifts of pyrene absorption maxima<sup>74</sup>

(v) Hybridization of X- or Y-modified ONs with cDNA leads to reduced fluorescence emission from the pyrene moieties,<sup>74,76</sup>



**Fig. 10** Side and top views of low energy structures for a duplex between an X-modified ON and cDNA (left) and a DNA duplex with a +1 zipper of X monomers (right). Reproduced from ref. 76 with permission from the Royal Society of Chemistry.



consistent with nucleobase-mediated quenching of emission from intercalating pyrenes<sup>84,85</sup>

(vi) The pyrene moiety of a duplex between an X-modified ON and cDNA remained intercalated throughout a 5 ns molecular dynamics simulation (Fig. 10, left),<sup>76</sup> consistent with earlier NMR studies suggesting 3'-directed intercalation of the pyrene moiety<sup>72</sup>

(vii) The destabilization of duplexes with +1 interstrand zippers of X monomers is also due to less favorable changes in enthalpy<sup>76</sup>

(viii) DNA duplexes with +1 zippers of X monomers exhibit blue-shifted pyrene absorption maxima relative to DNA duplexes with other zippers of X monomers or duplexes between X-modified ONs and cDNA, indicative of disrupted stacking interactions between pyrenes and nucleobases<sup>76</sup>

(ix) Steady-state fluorescence emission spectra of DNA duplexes with +1 interstrand zippers of X or Y monomers also display excimer signals at  $\lambda_{em} \sim 490$  nm, albeit less distinctly than the corresponding L-based Invader LNAs.<sup>62</sup> Interestingly, DNA duplexes with +2 interstrand zipper arrangements of X monomers, display a more prominent excimer signal, presumably due to pyrene-pyrene stacking in the major groove,<sup>62</sup> in a similar manner as observed for DNA duplexes with +2 zipper arrangements of pyrene-functionalized ara-uridine monomers.<sup>86-88</sup>

(x) Molecular dynamics simulations suggest that DNA duplexes with +1 zippers of X monomers are distorted. Similarly to the observations with Invader LNAs, two possible binding modes appear feasible, *i.e.*, one where both pyrenes intercalate into the duplex core (Fig. 10, right) and another where one pyrene intercalates while the other points into the major groove. The observed excimer emission suggests that forced intercalation of both pyrene moieties occurs (first mode), but the weak intensity of this emission suggests that the second mode is also adopted. Both binding modes can explain the labile nature of the double-stranded probes, as base pairing is distorted near the pyrene zipper in either duplex structure. Moreover, the presence of a large hydrophobic pyrene in the major groove (as in the second mode) likely perturbs the hydration spine of the duplex.<sup>76</sup>

**Key point:** Duplexes modified with 2'-*O*-(pyren-1-yl)methyl-uridine monomer X or 2'-*N*-methyl-2'-*N*-(pyren-1-yl)methyl-2'-aminodeoxyuridine monomer Y display very similar thermal

denaturation characteristics as corresponding duplexes modified with 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA-T monomer L. Biophysical characterization and molecular modeling simulations suggest that the pyrene moieties of the three monomers adopt similar binding modes.

### Preliminary dsDNA-recognition experiments using second-generation Invader probes

Having confirmed that the X and Y monomers display similar thermal denaturation properties and pyrene binding modes as monomer L, we proceeded to evaluate the dsDNA-targeting properties of Invader probes constructed using these monomers. Towards this end, a 3'-digoxigenin (DIG)-labeled DNA hairpin (DHP) – featuring a 9-mer double-stranded stem that is complementary to the Invader probes and linked at one end through a decameric thymidine ( $T_{10}$ ) loop – was used as a model target (Fig. 11).<sup>62</sup> The unimolecular and high-melting nature of the DNA hairpin ( $T_m = 58$  °C) ensures that both target strands are present in identical amounts and less likely to fray than a corresponding linear dsDNA target. Invasion of the double-stranded stem region is expected to result in the formation of a ternary invasion complex (IC) that has a lower mobility than the unbound DNA hairpin when incubation mixtures are resolved by nd-PAGE.

Importantly, L-, X-, and Y-based Invader probes were found to display near-identical recognition characteristics. Thus, 45–48% recognition of the DNA hairpin was observed when it was incubated with a 100-fold molar probe excess at room temperature (Fig. 12a).<sup>62</sup> However, an Invader probe constructed using the adenine analogue of monomer X, resulted in more efficient recognition of the hairpin target than the corresponding probe based on the adenine analogue of monomer L (41% *vs.* 17% recognition, 100-fold molar excess), possibly as the former is more energetically activated for dsDNA-recognition ( $T_A = 20.5$  °C *vs.* 18.5 °C, respectively) due to a slightly higher cDNA affinity.<sup>62</sup>

None of the following control experiments resulted in significant target recognition: (i) incubation of the DNA hairpin target with a 500-fold molar excess of a corresponding unmodified

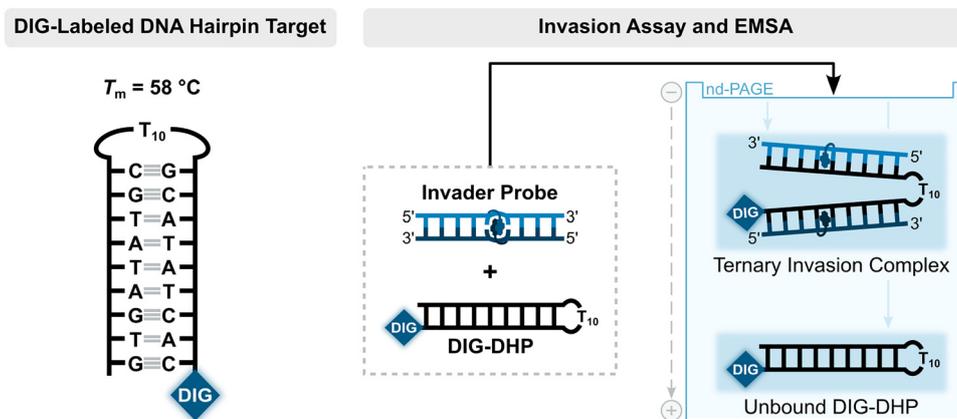


Fig. 11 Illustration of a model digoxigenin (DIG)-labeled DNA hairpin (DHP) target and the invasion assay used to assess dsDNA-recognition by Invader probes.



9-mer DNA duplex, (ii) incubation of the DNA hairpin target with a 100-fold molar excess of individual Invader strands, or (iii) incubation of fully base-paired DNA hairpins differing in sequence at one or two positions relative to Invader probes used at 100-fold excess (Fig. 12b).<sup>62</sup> Subsequent studies revealed that double-stranded probes with +5 or -3 interstrand zippers of X monomers also fail to recognize the corresponding DNA hairpin target – at conditions in which corresponding probes with a +1 zipper resulted in substantial recognition.<sup>89</sup> These results demonstrated that (i) the presence of energetic hotspots is indispensable for activating Invader probes for mixed-sequence dsDNA-recognition, (ii) both strands of an Invader probe are needed for efficient dsDNA-recognition, and (iii) recognition of dsDNA targets using Invader probes proceeds with excellent binding specificity.

**Key point:** 2'-O-(Pyren-1-yl)methyl-uridine monomer X and 2'-N-methyl-2'-N-(pyren-1-yl)methyl-2'-aminodeoxyuridine monomer Y are suitable building blocks for the construction of Invader probes, and display near-identical dsDNA-recognition characteristics as Invader probes based on 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA-T monomer L, including excellent binding specificity. Identification of these second-generation building blocks opened the possibility for optimization of probe designs and the use of Invader probes for exploratory diagnostic applications.

## Establishing design rules for second-generation Invader probes

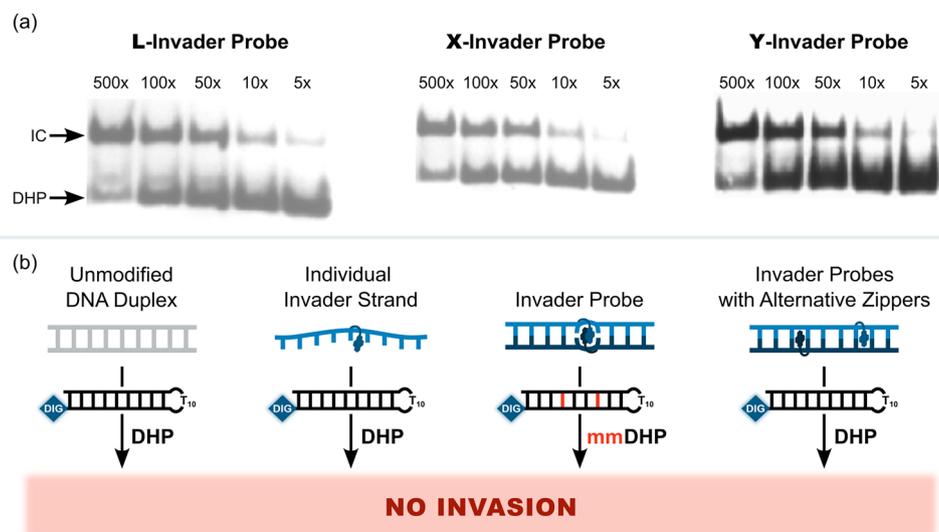
### The nucleobase composition of the energetic hotspot impacts the stability and dsDNA-recognition efficiency of Invader probes

With more easily accessible second-generation Invader monomers available, we were able to systematically study the impact that the nucleobase moieties of these monomers have on the

dsDNA-recognition efficiency of Invader probes. Towards this end, ten Invader probes were designed, differing only in the composition of a central energetic hotspot constructed using 2'-O-(pyren-1-yl)methyl-RNA monomers (Table 2). The inherent symmetry of the energetic hotspots, along with the utilized sequence context – *i.e.*, hotspots flanked by A:T pairs on either side – allowed us to study the sixteen possible hotspot compositions using only ten probes.

Each of these Invader probes was found to display reduced stability relative to the corresponding unmodified DNA duplexes with  $\Delta T_m$  values ranging between  $-17.5$  °C and  $0.0$  °C, and  $\Delta\Delta G_{293}$  values between +3 and +15 kJ mol<sup>-1</sup>, Table 2).<sup>79</sup> Probes with 2'-O-(pyren-1-yl)methyl-RNA-G monomers were particularly labile. Thus, +1 interstrand zippers of 2'-O-(pyren-1-yl)methyl-RNA monomers appear to disrupt stacking and nearby base pairs, irrespective of the hotspot's nucleobase composition. Further support for this was provided by the following characteristics that were observed for each Invader probe: (i) presence of pyrene excimer emission in steady-state fluorescence emission spectra, (ii) broad denaturation profiles indicating a distorted duplex, and (iii) blue-shifted pyrene absorption maxima ( $\Delta\lambda = 1-9$  nm) relative to duplexes between individual probe strands and cDNA.<sup>79</sup>

Duplexes between individual Invader strands modified with 2'-O-(pyren-1-yl)methyl-RNA-A/C/U monomers and cDNA were found to be strongly stabilized compared to the corresponding unmodified DNA duplexes ( $\Delta T_m$  values between +3.5 °C and +13.0 °C, and  $\Delta\Delta G^{293}$  values between  $-3$  kJ mol<sup>-1</sup> and  $-15$  kJ mol<sup>-1</sup>, Table 2).<sup>79</sup> In contrast, probe-target duplexes are less stable if modified with 2'-O-(pyren-1-yl)methyl-RNA-G monomers ( $\Delta T_m$  values between  $-5.5$  °C and +6.5 °C;  $\Delta\Delta G^{293}$  values between +6 kJ mol<sup>-1</sup> and  $-4$  kJ mol<sup>-1</sup>). Probe-target duplex stability is also impacted by the nature of the 3'-flanking



**Fig. 12** (a) Representative electrophoretograms from recognition experiments in which a DIG-DHP was incubated with varying probe excesses of L-, X-, or Y-modified probes. (b) Illustration depicting assay conditions that do not result in invasion of the DHP target. “IC” = invasion complex. “mmDHP” is a fully base-paired DNA hairpin that differs in sequence at two positions relative to the Invader probe. Data shown in panel a is from ref. 62, with permission from the American Chemical Society, copyright 2013.



Table 2  $\Delta T_m$ , TA,  $\Delta\Delta G^{293}$ , and  $\Delta G_{rec}^{293}$  values for Invader probes and duplexes between individual Invader strands and cDNA with varying hotspot compositions<sup>a</sup>

SEQUENCE				COLOR KEY				
5'-GTGA <b>BB</b> TGC		<b>B</b> = <b>A</b> / <b>G</b> / <b>C</b> / <b>U</b>		Less Favorable  More Favorable				
3'-CACT <b>BB</b> ACG		B = A / G / C / T						
Hotspot Composition	$\Delta T_m$ (°C)			TA (°C)	$\Delta\Delta G^{293}$ (kJ/mol)			$\Delta G_{rec}^{293}$ (kJ/mol)
	5'-Strand vs. cDNA	3'-Strand vs. cDNA	Invader Duplex		5'-Strand vs. cDNA	3'-Strand vs. cDNA	Invader Duplex	
5'- <b>AC</b> / 3'- <b>TG</b>	6.5	-1.0	-11.5	17.0	-5.0	0.0	9.0	-14.0
5'- <b>AG</b> / 3'- <b>TC</b>	8.5	5.0	-5.0	18.5	-15.0	-3.0	5.0	-23.0
5'- <b>AT</b> / 3'- <b>TA</b>	8.5	7.0	-6.5	22.0	-4.0	-4.0	5.0	-13.0
5'- <b>CA</b> / 3'- <b>GU</b>	13.0	10.0	0.0	23.0	-12.0	-10.0	4.0	-26.0
5'- <b>CC</b> / 3'- <b>GG</b>	6.5	5.5	-6.5	18.5	-5.0	-3.0	10.0	-18.0
5'- <b>CG</b> / 3'- <b>GC</b>	11.0	10.0	-2.0	23.0	-8.0	-9.0	7.0	-24.0
5'- <b>GA</b> / 3'- <b>CU</b>	6.5	3.5	-9.5	20.0	-4.0	-3.0	7.0	-14.0
5'- <b>GC</b> / 3'- <b>CG</b>	-3.5	-5.5	-17.5	8.5	4.0	6.0	15.0	-5.0
5'- <b>UA</b> / 3'- <b>AU</b>	13.0	13.0	-3.0	29.0	-14.0	-12.0	3.0	-29.0
5'- <b>UT</b> / 3'- <b>AA</b>	6.0	11.0	-5.0	22.0	-3.0	-11.0	5.0	-19.0

<sup>a</sup> **A**, **G**, **C**, and **U** are 2'-*O*-(pyren-1-yl)methyl-RNA monomers with adenin-9-yl, guanin-9-yl, cytosin-1-yl, and uracil-1-yl nucleobases, respectively. Hotspot compositions correspond to the sequence 5'-GTGA-**BB**-TGC/3'-CACT-**BB**-ACG where **B** and B are the indicated Invader and DNA monomers, respectively.

nucleotide relative to the 2'-*O*-(pyren-1-yl)methyl-RNA monomer, in a similar manner as observed with 2'-*N*-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA monomers.<sup>54</sup> Thus, greater stabilization is observed when 2'-*O*-(pyren-1-yl)methyl-RNA monomers are flanked by a purine on the 3'-side, consistent with 3'-directed pyrene intercalation.

As a result of the above factors, each of these Invader probes was found to exhibit favorable thermodynamics for the recognition of complementary dsDNA targets, albeit to varying degrees ( $\Delta G_{rec}^{293}$  values between  $-29$  and  $-5$  kJ mol<sup>-1</sup>, Table 2).<sup>79</sup> Invader probes with energetic hotspots composed exclusively of 2'-*O*-(pyren-1-yl)methyl-RNA pyrimidine monomers display the most prominent driving force for dsDNA-recognition ( $\Delta G_{rec}^{293}$  values between  $-29$  and  $-24$  kJ mol<sup>-1</sup>), which reflects the greater probe-target duplex stability when 2'-*O*-(pyren-1-yl)methyl-RNA monomers are flanked by 3'-purines.<sup>79</sup> In contrast, Invader probes with hotspots comprised exclusively of 2'-*O*-(pyren-1-yl)methyl-RNA-purine monomers – meaning each monomer is flanked by a 3'-pyrimidine – display the least favorable thermodynamics for dsDNA-recognition ( $\Delta G_{rec}^{293}$  values between  $-14$  and  $-5$  kJ mol<sup>-1</sup>).<sup>79</sup> Invader probes with mixed pyrimidine/purine hotspots, fall between these two extremes ( $\Delta G_{rec}^{293}$  values between  $-23$  and  $-14$  kJ mol<sup>-1</sup>).<sup>79</sup>

Four 9-mer Invader probes, each featuring a central hotspot comprised of 2'-*O*-(pyren-1-yl)methyl-RNA **U/U**, **C/C**, **U/A**, and **G/G** monomers (selected for their wide range of  $\Delta G_{rec}^{293}$  values and absolute  $T_m$  values), were evaluated using the DNA hairpin

invasion assay discussed previously (Fig. 11). Incubation of these Invader probes with the respective DNA hairpin targets at room temperature resulted in the formation of invasion complexes, in which the extent of invasion was consistent with the  $\Delta G_{rec}^{293}$  values of the probes. Thus, the Invader probe with the most favorable  $\Delta G_{rec}^{293}$  value (**U/U**) resulted in  $\sim 80\%$  recognition of the hairpin target when used at a 100-fold excess, while the probe with the least favorable  $\Delta G_{rec}^{293}$  value (**G/G**) resulted in minimal target recognition.

Thus, these results suggested to us that the preferred hotspot design entails the use of pyrene-functionalized pyrimidine monomers (targeting 5'-CA/CG/TA/TG-3' steps), although the corresponding adenine monomers are acceptable (targeting 5'-AA/AG/AT/CT/TT-3' steps); the use of guanine monomers is less desirable (avoid targeting 5'-AC/CC/GA/GC/GG/GT/TC steps) (Fig. 13).

The design principles inferred from the 9-mer study were successfully applied to longer and more densely modified Invader probes. Four 14-mer Invader probes, each featuring three distributed energetic hotspots, were designed to yield probes with different dsDNA-recognition potential. Thus, one of these Invader probes featured energetic hotspots solely composed of 2'-*O*-(pyren-1-yl)methyl-RNA pyrimidine monomers (efficient target recognition expected), while a different Invader probe only featured energetic hotspots composed of 2'-*O*-(pyren-1-yl)methyl-RNA purine monomers (least target recognition expected). Dose-dependent recognition of the corresponding



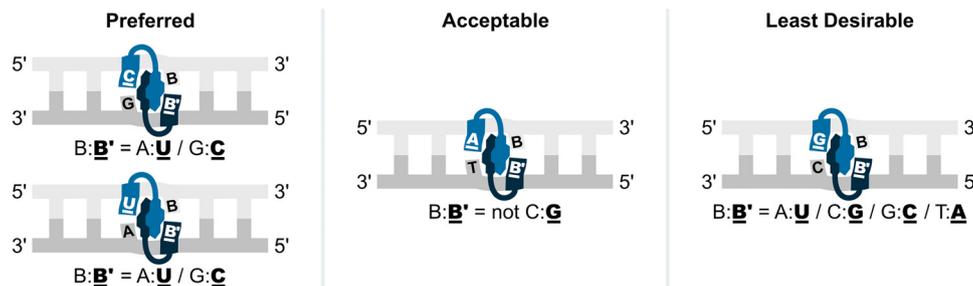


Fig. 13 Illustration of preferred, acceptable, and least desirable Invader hotspot compositions. B' and B are an Invader monomer and the complementary DNA monomer.

DNA hairpin target (GC content  $\sim 43\%$ ) was observed at room temperature with each of these Invader probes, with trends following expectations. Interestingly, even the Invader probe exclusively featuring hotspots composed of 2'-O-(pyren-1-yl)methyl-RNA-purine monomers, resulted in some target recognition, indicating that a favorable probe architecture (*i.e.*, densely modified probes, *vide infra*), can overcome sub-optimal hotspot selection.

**Key point:** Invader probes with energetic hotspots composed of 2'-O-(pyren-1-yl)methyl-RNA pyrimidine monomers display the most prominent dsDNA-recognition potential due to the high stability of the resulting probe-target duplexes where purines flank the monomers on the 3'-side. Conversely, 2'-O-(pyren-1-yl)methyl-RNA-G monomers are best avoided as individual probe strands display low cDNA affinity when cytidines flank the monomers on the 3'-side.

### Extensively modified second-generation Invader probes display improved dsDNA-recognition characteristics

Invader probes with different modification patterns (*i.e.*, variable number of, location of, and distance between energetic hotspots) were constructed using second-generation monomers in a preferred, TA-alternating sequence context and studied to understand how probe designs impact dsDNA-recognition.<sup>90</sup> Twenty 13-mer Invader probes, modified with X or Y monomers were synthesized toward this end (Table 3).

As in the preliminary studies, singly modified ONs formed greatly stabilized duplexes with cDNA relative to unmodified ONs, with Y-modified ONs consistently forming slightly more stable duplexes ( $\Delta T_m = 7.0\text{--}11.0\text{ }^\circ\text{C}$  vs.  $8.0\text{--}13.5\text{ }^\circ\text{C}$  for X- and Y-modified duplexes, respectively).<sup>90</sup> Incorporation of a second pyrene-functionalized monomer resulted in further stabilization ( $\Delta T_m = 14.0\text{--}21.5\text{ }^\circ\text{C}$ ) although the observed  $T_m$  increases were slightly less than additive, suggesting that intercalation of a first pyrene moiety, to a minor degree, impacts the stabilizing properties of a second pyrene moiety. This trend becomes progressively more pronounced with increasing modification levels.

Invader probes with a single energetic hotspot were found to display similar  $T_m$ s as the corresponding unmodified DNA duplexes ( $\Delta T_m = -4.0$  to  $+3.0\text{ }^\circ\text{C}$ ),<sup>90</sup> while probes with two or four hotspots were moderately stabilized ( $\Delta T_m = +2.0$  to  $+12.5\text{ }^\circ\text{C}$ ), suggesting that the presence of a first energetic

Table 3 TA and  $\Delta G_{rec}^{293}$  values for different X- and Y-modified probe designs

ON	Sequence	M = Monomer X		M = Monomer Y	
		TA ( $^\circ\text{C}$ )	$\Delta G_{rec}^{293}$ (kJ/mol)	TA ( $^\circ\text{C}$ )	$\Delta G_{rec}^{293}$ (kJ/mol)
M1	5'-GGMATATATAGGC	18.0	-27	22.0	-35
M2	3'-CCAMATATATCCG				
M3	5'-GGTAMATATAGGC	22.0	-32	21.5	-38
M4	3'-CCATAMATATCCG				
M5	5'-GGTATAMATAGGC	22.0	-32	22.5	-39
M6	3'-CCATATAMATCCG				
M7	5'-GGTATATAMAGGC	21.0	-31	22.0	-41
M8	3'-CCATATATAMCCG				
M9	5'-GGMAMATATAGGC	29.5	-40	25.0	-63
M10	3'-CCAMAMATATCCG				
M11	5'-GGMATAMATAGGC	23.5	-46	29.0	-58
M12	3'-CCAMATAMATCCG				
M13	5'-GGMATATAMAGGC	21.5	-38	32.0	-69
M14	3'-CCAMATATAMCCG				
M15	5'-GGTAMAMATAGGC	28.5	-45	35.0	-67
M16	3'-CCAYAMAMATCCG				
M17	5'-GGTATAMAMAGGC	24.0	-46	28.0	-49
M18	3'-CCATATAMAMCCG				
M19	5'-GGMAMAMAMAGGC	45.5	-71	56.5	-95
M20	3'-CCAMAMAMAMCCG				

hotspot reduces the destabilizing characteristics of subsequent hotspots. Lower  $T_m$ s were generally observed for Y-modified probes and probes with consecutive energetic hotspots, mirroring observations for Invader probes modified with the original 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA L monomers in the same sequence context.<sup>63</sup>

The most densely modified Invader probes are most strongly activated for recognition of isosequential dsDNA targets (TA values =  $18.0\text{--}22.5\text{ }^\circ\text{C}$ ,  $21.5\text{--}35.0\text{ }^\circ\text{C}$ , and  $45.5\text{--}56.5\text{ }^\circ\text{C}$ , for Invader probes with one, two, and four hotspots, respectively, Table 3), even though "TA per hotspot" values decrease with increasing levels of modification.<sup>90</sup> Similar conclusions are reached by observing changes in free energy upon duplex formation. Thus, the Invader probes were found to be strongly activated for recognition of isosequential dsDNA targets (*i.e.*,  $\Delta G_{rec}^{293}$  values  $\ll 0\text{ kJ mol}^{-1}$ , Table 3) due to the low stability of the double-stranded probes (*i.e.*,  $\Delta\Delta G^{293}$  values between  $-6\text{ kJ mol}^{-1}$  and  $+11\text{ kJ mol}^{-1}$ , relative to the corresponding dsDNA targets) and high stability of the probe-target duplexes (*i.e.*,  $\Delta\Delta G^{293}$  values between  $-52\text{ kJ mol}^{-1}$  and  $-6\text{ kJ mol}^{-1}$ ). Invader probes with multiple energetic hotspots displayed



more favorable thermodynamics for dsDNA-recognition than single hotspot probes, and Y-modified Invader probes were more strongly activated for dsDNA-recognition than the corresponding X-modified probes ( $\Delta G_{\text{rec}}^{293}$  values more favorable by 3–31 kJ mol<sup>-1</sup>, Table 3). Interestingly, the X- and Y-modified Invader probes displayed more favorable thermodynamics for dsDNA-recognition than the corresponding probes based on the original<sup>63</sup> 2'-N-(pyren-1-yl)methyl-2'-amino- $\alpha$ -L-LNA L monomers ( $\Delta G_{\text{rec}}^{293}$  values more favorable by 1–29 kJ mol<sup>-1</sup>), further highlighting their potential as second-generation monomers. Recognition of dsDNA by these Invader probes is very strongly enthalpically favored ( $\Delta H_{\text{rec}} \ll 0$  kJ mol<sup>-1</sup>), further pointing to forced intercalation as the main driving force (stabilizing in probe-target duplexes and destabilizing in Invader probes).

Each of the twenty Invader probes recognized the corresponding DNA hairpin target to some degree (20–100% recognition) when incubated at 200-fold molar excess at room temperature for 17 h. Six of the probes – featuring one, two, or four consecutive hotspots of X or Y monomers – were evaluated in more detail in dose-response experiments to determine  $C_{50}$  values, *i.e.*, the probe concentration resulting in 50% recognition of the DNA hairpin targets (Fig. 14a). Increasing the hotspot density progressively decreased  $C_{50}$  values from single digit micromolar to submicromolar ranges. Y-based probes displayed lower  $C_{50}$  values than X-based probes in agreement with the observed  $\Delta G_{\text{rec}}^{293}$  trends.

Room temperature incubation of the DNA hairpin target with a 1000-fold excess of individual probe strands with a single X or Y incorporation, or a 500-fold excess of probe strands with two X or Y incorporations, resulted in no or merely trace formation of recognition complexes after 17 hours, whereas incubation with the corresponding Invader probes resulted in complete recognition (Fig. 14c). Conversely, room temperature incubation with a 50-fold excess of individual probe strands with four incorporations of X or Y monomers generally resulted in complete recognition of the DNA hairpin (Fig. 14c).<sup>90</sup>

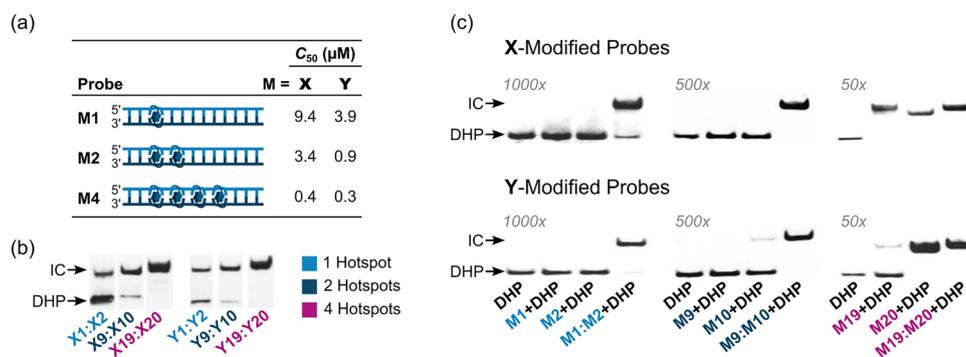
However, only trace recognition was observed when a 2-fold excess of the individual strands with four X modifications were incubated at 37 °C, *i.e.*, conditions resulting in complete recognition of the DNA hairpin target with the corresponding Invader probe.<sup>37</sup> These observations demonstrate that while the high cDNA affinity of densely modified individual Invader strands can facilitate recognition of DNA duplexes under certain conditions, the corresponding Invader probes result in more efficient recognition.

A subsequent study sought to investigate design parameters governing target recognition in other, non-TA-rich sequence contexts.<sup>91</sup> Ten Cy3-labeled Invader probes were designed varying in the number of 2'-O-(pyren-1-yl)methyl-RNA A/C/U monomers (20–30% modified), length (14–16 base pairs), and GC-content (30–70%) (Table 4).

Neither the modification density, length, or GC-content of the probes were found to be predictive of Invader probe lability as neither parameter correlated with absolute ( $T_m$  or  $\Delta G^{310}$ ) or relative ( $\Delta T_m$  or  $\Delta \Delta G^{310}$ ) measures of Invader probe stability. In contrast, the modification density correlated with  $\Delta T_m$  or  $\Delta \Delta G^{310}$  values for probe-cDNA duplexes, highlighting this design parameter as a predictor of high probe-target duplex stability ( $\Delta T_m$  values between +3.5 °C and +22.0 °C in this study).

As a consequence of the low stability of the probes ( $\Delta \Delta G_{310}$  values between -1 kJ mol<sup>-1</sup> and +25 kJ mol<sup>-1</sup>) and high stability of the probe-target duplexes ( $\Delta \Delta G_{310}$  values between -33 kJ mol<sup>-1</sup> and +19 kJ mol<sup>-1</sup>), every Invader probe was found to be thermodynamically activated for dsDNA-recognition ( $\Delta G_{\text{rec}}^{310}$  values between -56 kJ mol<sup>-1</sup> and -7 kJ mol<sup>-1</sup>;  $T_A$  values between +13.5 °C and +44.5 °C, Table 4).

Four of the ten Invader probes resulted in efficient recognition of the corresponding DNA hairpin targets when incubated for ~15 hours at 37 °C ( $C_{50}$  values between 0.2  $\mu$ M and 0.7  $\mu$ M), five probes displayed moderately efficient recognition ( $C_{50}$  ~ 1–10  $\mu$ M), while one probe failed to recognize the DNA hairpin target (Table 4).



**Fig. 14** Invasion of DNA hairpins by 13-mer Invader probes featuring one, two, or four X- or Y-based energetic hotspots. (a)  $C_{50}$  values for X- and Y-modified Invader probes varying in hotspot density. (b) Representative electrophoretograms for dsDNA-recognition experiments in which the indicated probes were incubated at 200-fold molar excess with the corresponding DHP target. (c) DHP invasion experiments using individual Invader strands and the corresponding Invader probes at 1000-, 500-, or 50-fold molar excess, as indicated. "IC" = invasion complex. Data and figures adapted from ref. 90 with permission from the Royal Society of Chemistry.



Table 4 Sequences and key parameters for Cy3-labeled Invader probes modified with 2'-O-(pyren-1-yl)methyl-RNA A/C/U monomers<sup>a</sup>

Probe	Sequence	Modification density	TA (°C)	$\Delta G_{\text{rec}}^{293}$ (kJ/mol)	$C_{50}$ ( $\mu\text{M}$ )	Approximate nuclei detected (%)	
						d-FISH	nd-FISH
INV1	5'-Cy3-TUATCAGCACUGUGC-3' 3'-AAUAGTCGTGACACG-Cy3-5'	20.0%	+23.5	ND	1.3	0	0
INV2	5'-Cy3-AUACUGGTTTUGUTC-3' 3'-TAUGACCAAACACAAG-Cy3-5'	25.0%	+44.5	-50	0.2	90	85
INV3	5'-Cy3-TUGUGCCCTGGCAAC-3' 3'-AACACGGGACCGTUG-Cy3-5'	20.0%	ND	ND	2.9	40	30
INV4	5'-Cy3-AGCCCUGTGCCCTG-3' 3'-TCGGGACACGGGAC-Cy3-5'	21.4%	+23.0	-7	ND	90	90
INV5	5'-Cy3-GATTTTCAGCCAUGUGC-3' 3'-CTAAGTCGGTACACG-Cy3-5'	18.8%	+30.5	-29	4.1	0	0
INV6	5'-Cy3-CUGUGCAACTGGTUTG-3' 3'-GACACGTTGACCAAC-Cy3-5'	18.8%	+13.5	-22	>10	60	20
INV7	5'-Cy3-CUGUGCAAUATTTUGT-3' 3'-GACACGTTATAAAACA-Cy3-5'	25.0%	+38.0	-46	0.7	60	25
INV8	5'-Cy3-TTCACAGCCCUGUGC-3' 3'-AAGUGTCGGGACACG-Cy3-5'	20.0%	+26.5	-52	0.6	60	25
INV9	5'-Cy3-TUAUATGCTGUTCTC-3' 3'-AAUAUACGACAAGAG-Cy3-5'	20.0%	+21.5	-19	1.5	60	0
INV10	5'-Cy3-GUGUAGTGUUATG-3' 3'-CACAUACAUUAC-Cy3-5'	28.6%	+40.5	-56	0.2	90	90
OPT6	5'-Cy3-CUGUGCAACUGGTUTG-3' 3'-GACACGUTGACCAAC-Cy3-5'	31.3%	+43.0	-84	>10	90	85
OPT8	5'-Cy3-TTCACAGCCCUGUGC-3' 3'-AAGUGUCGGGACACG-Cy3-5'	26.7%	+55.5	-93	0.2	90	75
OPT9	5'-Cy3-TUAUAGCUGUTCTC-3' 3'-AAUAUACGACAAGAG-Cy3-5'	33.3%	+54.5	-59	0.6	75	25

<sup>a</sup> ND = not determined.

TA/ $\Delta G_{\text{rec}}^{310}/C_{50}$  values correlated with the hotspot densities of the probes, further highlighting the latter as the key parameter for the design of efficient dsDNA-targeting Invader probes (Fig. 15). This insight was used to improve the dsDNA-targeting properties of suboptimal Invader probes. Thus, increasing the hotspot density of the probes markedly increased their TA values and lowered their  $C_{50}$  values to the low submicromolar range (compare INV6/8/9 and OPT6/8/9). The hotspot density was also found to be the key factor in determining the efficiency of Invader probes in fluorescence *in situ* hybridization (FISH) assays conducted under denaturing and non-denaturing conditions (*vide infra*).<sup>91</sup>

**Key point:** Invader probes with a density of second-generation monomers above 25% result in particularly efficient recognition of model dsDNA targets. This is due to increased cDNA affinity of individual probe strands with increasing modification densities, rather than additional destabilization of Invader probes. Recognition does not appear to be materially limited by the GC-content of a target region as recognition of model dsDNA targets with GC-contents of up to ~70% has been demonstrated. However, since the hotspots of Invader probes are preferably constructed using pyrene-functionalized pyrimidine monomers, this calls for dsDNA target regions with a high density of 5'-pyrimidine-purine steps (Fig. 13).



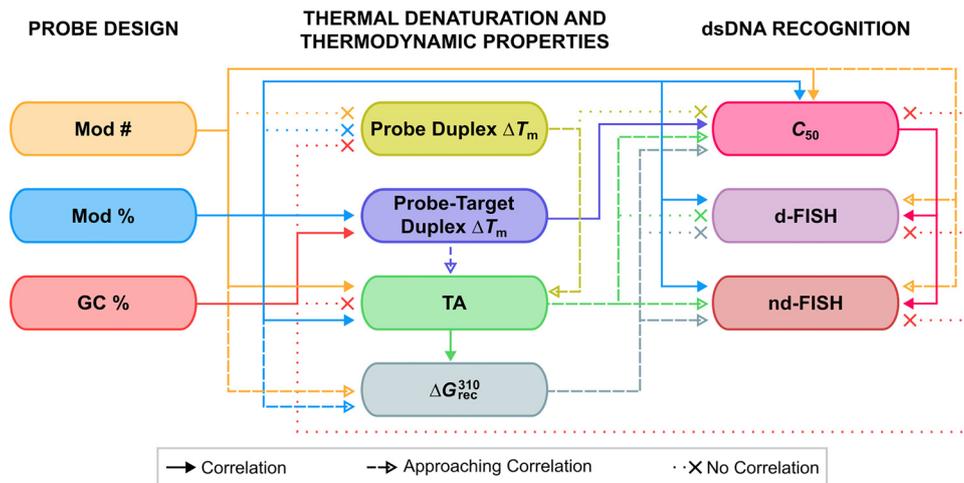


Fig. 15 Correlations between design parameters governing recognition of dsDNA by Invader probes.

## Recognition kinetics, binding specificity and invasion mechanism of second-generation Invader probes

### Recognition kinetics of second-generation Invader probes

The kinetics of dsDNA-recognition were studied using 13-mer Invader probes featuring one, two, or four consecutive X- or Y-based hotspots (for sequences, see Table 3). Invader probes with two hotspots resulted in 50% invasion of the corresponding DNA hairpin target after  $\sim 3$  hours at room temperature when used at 200-fold excess, whereas 50% recognition was observed after only 10–50 min for Invader probes with four consecutive hotspots.<sup>90</sup> The corresponding pseudo-first order rate constants  $k_{obs}$  for the initial phases of DNA hairpin recognition of  $1.1 \times 10^{-3}/7.5 \times 10^{-3}/9.7 \times 10^{-2} \text{ min}^{-1}$  and  $1.4 \times 10^{-3}/4.3 \times 10^{-3}/1.5 \times 10^{-2} \text{ min}^{-1}$  were observed for the X- or Y-modified M1:M2, M9:M10 and M19:M20 Invader probes, respectively. Thus, increasing the modification density in second-generation Invader probes not only results in a greater degree of dsDNA-recognition but also faster recognition.

The extent and kinetics of Invader-mediated dsDNA-recognition are also strongly dependent on the experimental temperature.<sup>90</sup> For example, a 13-mer Invader probe with two consecutive hotspots of X monomers (X9:X10, see Table 3) failed to recognize the DNA hairpin target at 8 °C, whilst resulting in  $\sim 20\%$ ,  $\sim 50\%$ , and  $\sim 80\%$  recognition when incubated for only 10 min at 200-fold excess at room temperature, 35 °C, and 45 °C, respectively. Similar results were observed for corresponding Y-modified Invader probes. The temperature-dependent rate enhancements are mainly attributed to increased probe denaturation ( $T_m = 40.0$  °C) rather than partial DNA hairpin denaturation ( $T_m = 58.5$  °C). Incubation at 55 °C resulted in rapid, but incomplete DNA hairpin invasion, presumably as the invasion complex undergoes partial denaturation ( $T_m$ s for probe-target duplexes were between 51.5 and 55.5 °C).<sup>90</sup> The impact of incubation temperature and hotspot density on recognition kinetics was further underscored in

subsequent experiments. Thus, Invader probes with three X-based energetic hotspots were shown to recognize corresponding DNA hairpin targets with  $C_{50}$  values of 40–50 nM when incubated at 37 °C for 2.5 hours,<sup>92</sup> representing a  $\sim 10$ -fold improvement relative to Invader probes with four hotspots targeting the same DNA hairpin at room temperature for 17 hours.<sup>90</sup>

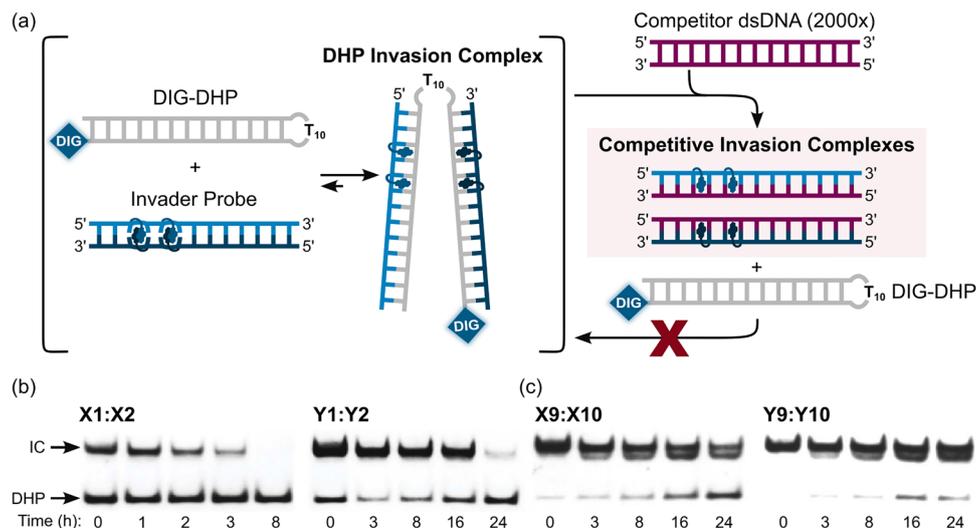
Once formed, invasion complexes entailing highly modified Invader probes are stable. Thus, 60–85% of invasion complexes – formed following incubation of a DNA hairpin target with Invader probes with two X- or Y-based hotspots – remained intact for 24 hours following treatment with a large excess of a competitor target (Fig. 16). Conversely, invasion complexes entailing single hotspot Invader probes dissociated faster (Fig. 16, only trace amounts observed after 8 h and 24 h for X1:X2 and Y1:Y2, respectively).

**Key point:** Faster invasion of dsDNA targets is observed with more extensively modified Invader probes and the resulting invasion complexes do not readily dissociate once formed. Invasion of dsDNA regions can be accelerated by raising incubation temperatures closer to or above the denaturation temperatures of the Invader probes but below the denaturation temperature of the invasion complexes.

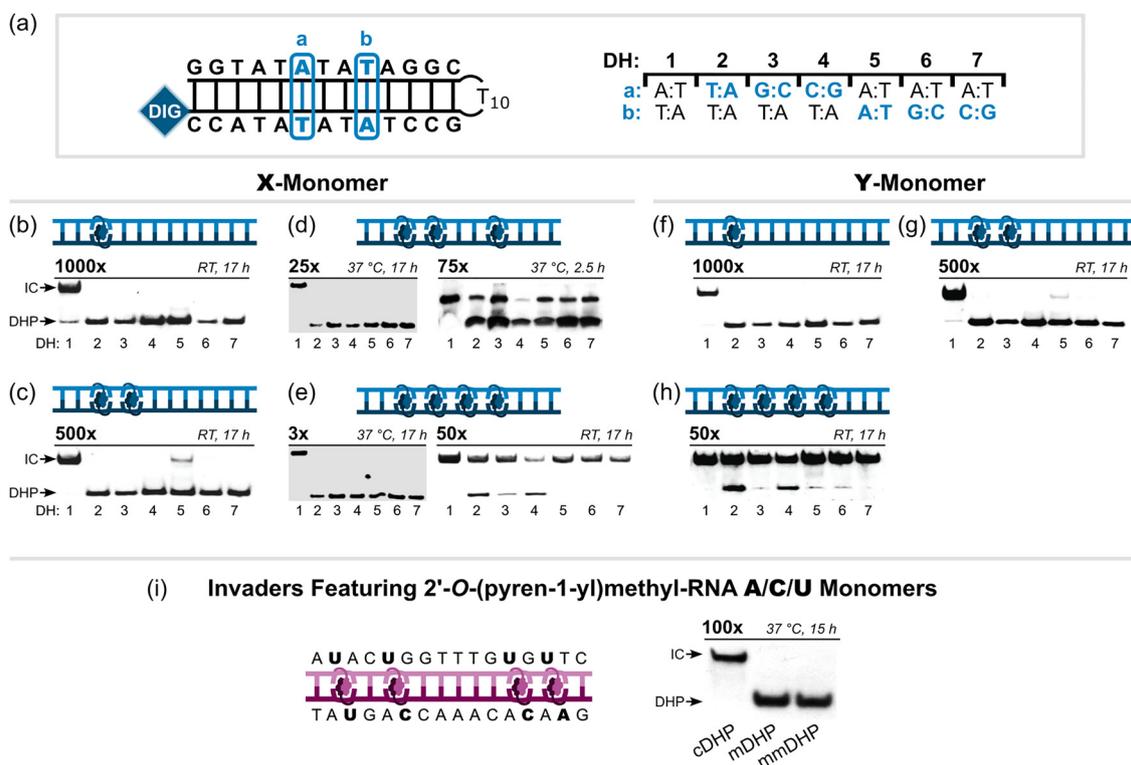
### Invader probes bind dsDNA regions with high specificity

In addition to having prominent affinity towards complementary dsDNA regions, Invader probes display outstanding binding specificity. Early studies demonstrated that 9-mer Invader probes with one hotspot based on first- or second-generation monomers fully discriminate DNA hairpins with stems differing in sequence at merely one or two positions relative to the probes.<sup>62</sup> Subsequently, 13-mer Invader probes with one or two hotspots of X or Y monomers were shown to display near-perfect discrimination of six DNA hairpins with stem regions differing only at a single central position at conditions resulting in near-complete recognition of the complementary DNA hairpins (Fig. 17b, c and f, g, respectively).<sup>90</sup> The corresponding





**Fig. 16** Dissociation kinetics of DHP invasion complexes. (a) Illustration of competition assay entailing pre-formed DHP-Invader invasion complexes and a large excess of competitor dsDNA target. (b) and (c) Representative electrophoretograms depicting 200-fold molar excess of Invader probe (see Table 3 for sequences) incubated with complementary DIG-DHP for 17 h at room temperature, followed by ( $t = 0$  h) addition of 2000-fold molar excess of a complementary linear dsDNA target. IC = invasion complex. Electrophoretograms in (b) and (c) were reproduced from ref. 90 with permission from the Royal Society of Chemistry.



**Fig. 17** Binding specificity of second-generation Invader probes. (a) Illustration of single-mismatch hairpin targets for experiments shown in b–h where “a” and “b” indicate the position (left side of panel) and sequence (right side of panel) of the mismatches. Specificity experiments entailing (b)–(e) X-based or (f)–(h) Y-based Invader probes with one to four hotspots. (i) Specificity experiments entailing a 16-mer mixed-base Invader probe with four interspersed hotspots constructed using 2'-O-(pyren-1-yl)methyl-RNA A/C/U monomers incubated with complementary, single-mismatch (position 8, T:A → A:T), and doubly-mismatch (position 5, T:A → A:T, and position 12, T:A → A:T) DNA hairpin targets (cDHP, mDHP, and mmDHP, respectively). Probe excess and assay conditions are indicated above each electrophoretogram. All electrophoretograms in this figure were reproduced from ref. 36, 37, and 90 with permission from the Royal Society of Chemistry.



Invader probes with four X- or Y-based hotspots resulted in near-complete recognition of the six non-target DNA hairpins when incubated at 50-fold excess for 17 hours at room temperature (Fig. 17e and h, respectively).<sup>90</sup> However, perfect discrimination of the non-target DNA hairpins was restored when the X-based Invader probe was used at 3-fold excess and an incubation temperature of 37 °C, *i.e.*, conditions resulting in complete recognition of the complementary DNA hairpin target (Fig. 17e).<sup>37</sup> Similarly, Invader probes with three X-based hotspots in the same sequence context also resulted in perfect discrimination of these non-target hairpins when used at 25-fold (17 h incubation, Fig. 17d)<sup>37</sup> or 5-fold excess (2.5 h incubation) at 37 °C,<sup>92</sup> whilst resulting in complete recognition of the complementary DNA hairpin. Partial recognition of non-target DNA hairpins was observed when this Invader probe was used at a very large 75-fold excess (37 °C, 2.5 h, Fig. 17d).<sup>92</sup> Remarkably, 16-mer mixed-base Invader probes with four interspersed hotspots based on 2'-O-(pyren-1-yl)methyl-RNA A/C/U monomers were found to discriminate DNA hairpins with stems that differ in sequence at only one or two positions relative to the probes when used at 100-fold excess (37 °C, 15 h), while resulting in complete recognition of the complementary targets (Fig. 17i).<sup>36</sup> Thus, high-affinity Invader probes can completely discriminate against dsDNA regions with ~94% sequence homology, which is expected to enable single nucleotide polymorphism (SNP)-specific recognition of dsDNA target regions.

The excellent dsDNA-binding specificity of Invader probes modified with second-generation monomers contrasts the moderate binding fidelity of individual singly X- or Y-modified strands. Thus, duplexes between centrally X- or Y-modified 9-mer ONs and DNA strands with mismatched C/G/T residues opposite of the pyrene-functionalized uridine monomers displayed  $\Delta T_m$  values of  $-13.0/-5.0/-6.5$  °C and  $-22.0/-3.5/-12.0$  °C relative to the corresponding matched duplexes, respectively, as compared to  $\Delta T_m$  values of  $-16.5/-9.5/-17.0$  °C for the corresponding unmodified ON.<sup>74</sup> Interestingly, ONs with two X- or Y-monomers, placed as next-nearest neighbors (*i.e.*, a favorable probe design for Invader probes), discriminate DNA strands with mismatched A/C/G nucleotides opposite of the XAX/YAY modification cassettes with improved fidelity, *i.e.*,  $\Delta T_m$  values of  $-24.0/-17.0/-14.0$  °C and  $-21.5/-10.5/-13.0$  °C compared to  $\Delta T_m$  values of  $-17.0/-15.5/-9.0$  for the corresponding unmodified ON.<sup>74</sup>

However, the high binding specificity of Invader probes likely arises primarily from stringency clamping effects, which are characteristic of structured, metastable probes.<sup>93</sup> Stringency clamping increases the free-energy difference between matched and mismatched complexes by employing an oligonucleotide with a secondary structure that is less stable than the recognition complex formed with the correct target, yet more stable than the corresponding complex formed with a mismatched target. As a result, the secondary structure of the probes interferes selectively with mismatched binding, yielding higher specificity than unstructured oligonucleotides.

Established examples of stringency clamping include hairpin-based probes (*e.g.*, molecular beacons) in duplex

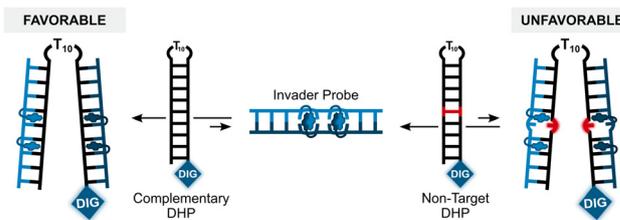


Fig. 18 Stringency clamping effects underlying the excellent binding specificity of Invader probes. The non-target DHP represents a DHP with a fully base-paired stem region but differing in sequence at one position (red) relative to the Invader probe.

formation and imperfect duplexes in triplex recognition.<sup>93</sup> By analogy, productive binding of Invader probes to non-target dsDNA would require both denaturation of the double-stranded probe and target region and formation of two mismatched duplexes within the invasion complex. The double-stranded Invader probes are presumed to be less stable than the ternary invasion complexes formed with the correct DNA hairpin target, but more stable than complexes formed with non-target sequences containing two destabilized mismatched duplexes (Fig. 18). Consequently, Invader probes function as stringency clamps by selectively disfavoring non-target binding and amplifying the thermodynamic discrimination between target and non-target interactions.

**Key point:** Invader probes based on second-generation monomers exhibit exceptional binding specificity due to stringency clamping effects and can discriminate against dsDNA regions with ~94% sequence homology.

#### Proposed mechanism for Invader probe-mediated dsDNA recognition

While no definitive proof for the mechanism of Invader-mediated dsDNA-recognition has been obtained, we stipulate that densely modified Invader probes are partially denatured and significantly perturbed, thereby enabling the pyrene moieties to initially interact non-specifically with – and intercalate into – particularly accessible regions of DNA duplexes. This may initiate nucleation and unzipping of Watson–Crick base pairs, ultimately resulting in recognition of complementary dsDNA regions *via* a double-duplex invasion mechanism that resembles the mechanism proposed for pseudo-complementary PNA (Fig. 19).<sup>31</sup>

## Comparison of second-generation Invader probes with other DNA-targeting probes

Having established the main design guidelines for second-generation Invader probes, we set out to compare their dsDNA-recognition characteristics with other strand-invading probes, including INAs and single-stranded <sup>Ser</sup>γPNAs, <sup>MP</sup>γPNAs, and partially or fully LNA-modified ONs.<sup>36,37,92</sup>



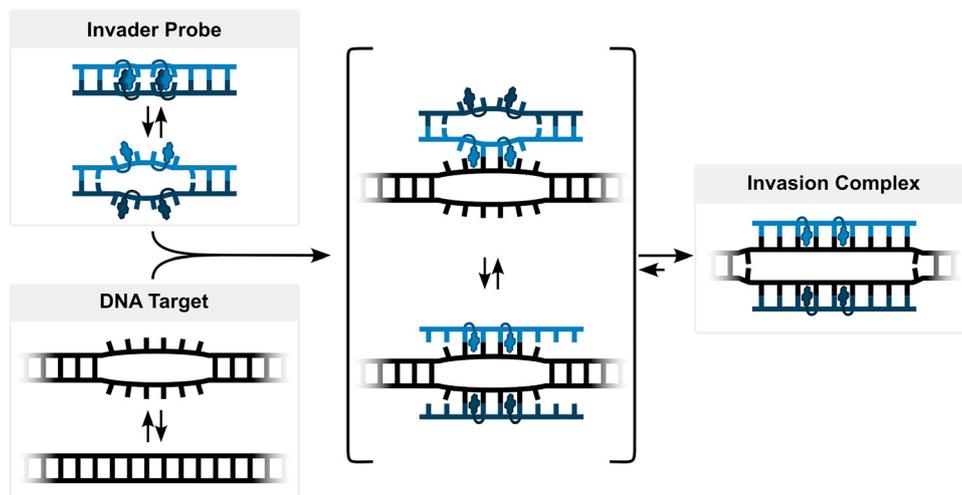


Fig. 19 Proposed double-duplex invasion mechanism for recognition of complementary dsDNA regions by Invader probes.

In a first comparison study,<sup>92</sup> two 13-mer Invader probes with three X-based hotspots recognized the corresponding DNA hairpin target with similar efficiency as the equivalent INA probes ( $C_{50} = 30\text{--}50\text{ nM}$ ,  $37\text{ }^{\circ}\text{C}$ , 2.5 h incubation), whereas  $\sim 30\%$  LNA-modified single-stranded ONs only resulted in

moderate dsDNA-recognition despite the affinity-enhancing properties of LNA monomers ( $\Delta T_m$  values of  $\sim 17\text{ }^{\circ}\text{C}$  were observed for duplexes with cDNA) (Fig. 20; for monomer structures see Fig. 2). Two  $^{MP}\gamma\text{PNA}$  probes were found to form very stable duplexes with cDNA ( $\Delta T_m > 40\text{ }^{\circ}\text{C}$ ), but just one

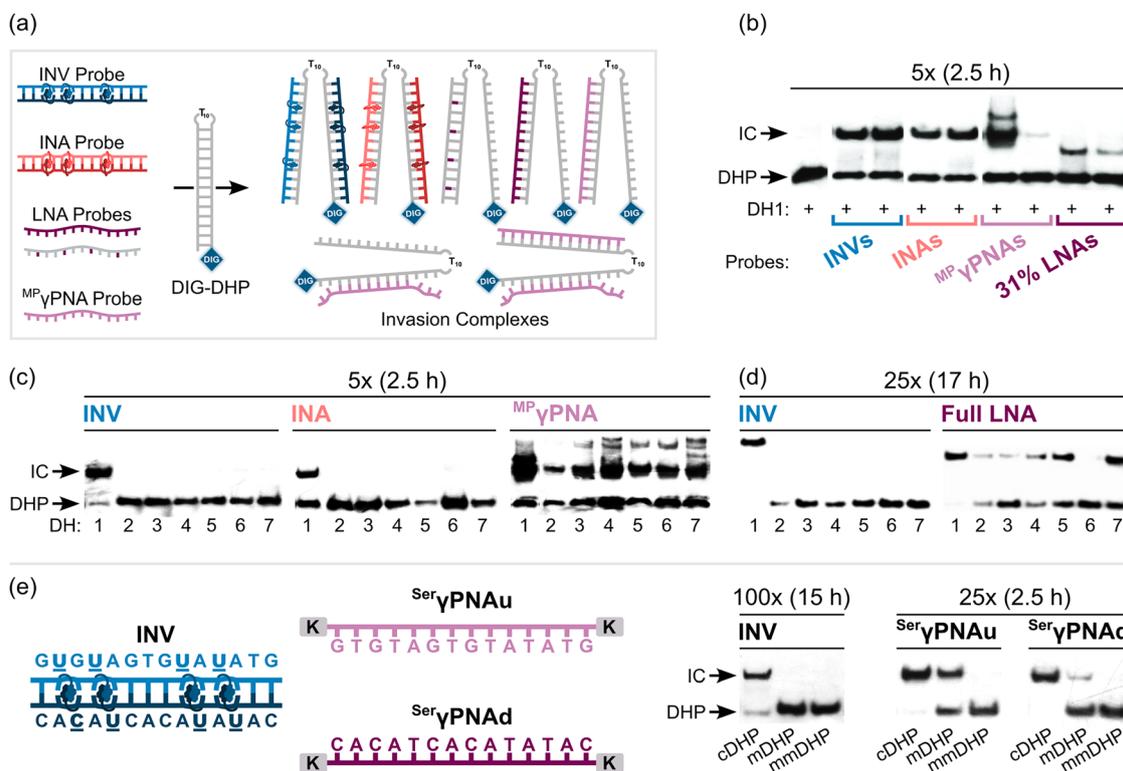


Fig. 20 Comparison of dsDNA-invasion characteristics of Invader, INA,  $^{MP}\gamma\text{PNA}$ , LNA, and  $^{Ser}\gamma\text{PNA}$  probes. (a) Illustration of the DIG-DHP invasion assay using various probes. (b) Representative gel electrophoretogram from dsDNA-invasion experiments in which **DH1** (complementary target to the probes) was incubated with the indicated probe type. (c) and (d) Representative gel electrophoretograms from binding specificity experiments. (e) Sequences of 14-mer Invader and  $^{Ser}\gamma\text{PNA}$  probes (K = lysine), and invasion assay results with complementary DHP (cDHP), singly mismatch (m), or doubly mismatch (mm) DHP targets. Probe excess and incubation times are as listed; all incubations were performed at  $37\text{ }^{\circ}\text{C}$ . For structures of modifications, see Fig. 2. See Fig. 17a for sequences of **DH1–DH7**. All electrophoretograms in this figure were reproduced from ref. 37 and 92 with permission from the Royal Society of Chemistry.



resulted in moderate recognition of the DNA hairpin ( $C_{50} \sim 250$  nM), whereas the other only resulted in trace recognition (Fig. 20). Both  $^{MP}\gamma$ PNA probes displayed a proclivity for secondary structure formation in the AT-rich, partially self-complementary sequence context studied (see Table 3 for sequence context) as concentration-dependent  $T_m$  trends – consistent with dimer formation – were observed in the absence of cDNA targets. Formation of stable secondary structures is commonly observed for single-stranded probes with affinity-increasing modifications, which may interfere with dsDNA-recognition.<sup>41</sup> Moreover, the dsDNA-binding  $^{MP}\gamma$ PNA probe was found to form a partial duplex with the isosequential (*i.e.*, non-target) arm of the DNA hairpin, indicating potential specificity challenges.

The double-stranded Invader and INA probes resulted in slower but far more specific dsDNA-recognition than the  $^{MP}\gamma$ PNA probe. Thus, when used at 37 °C and 5-fold excess relative to the DNA hairpin target, the  $^{MP}\gamma$ PNA probe resulted in 50% invasion in less than 10 min, while it took 37–68 min and 55–75 min, respectively for 50% invasion to be reached with the equivalent triple hotspot Invader and INA probes. This likely reflects the different recognition mechanisms of  $^{MP}\gamma$ PNA *vis-à-vis* Invader/INA probes, *i.e.*, single- *vs.* double-duplex invasion, respectively, with the latter requiring partial denaturation of the double-stranded probes to reveal nucleation sites that can initiate formation of invasion complexes with the dsDNA target (Fig. 19). Invader and INA probes resulted in complete discrimination of DNA hairpins with stems differing in sequence at one of thirteen positions *vis-à-vis* the probes, whereas the  $^{MP}\gamma$ PNA resulted in minimal discrimination of these targets (Fig. 20c), which highlights the advantage of the stringency clamping effect displayed by double-stranded probes.

A subsequent comparison study<sup>37</sup> – using the same DNA hairpins – revealed that the corresponding, fully modified LNA probes display similar dsDNA-recognition characteristics as the  $^{MP}\gamma$ PNA probes; one LNA probe resulted in moderate recognition ( $C_{50} \sim 300$  nM, 37 °C, 17 h incubation), while the other LNA probe only resulted in trace recognition. In contrast, Invader probes with three or four energetic hotspots displayed  $C_{50}$  values of  $\sim 85$  nM and  $\sim 35$  nM, respectively. The single-stranded LNA probe displayed poor discrimination of the non-target DNA hairpins when used at 25-fold excess, whereas the Invader probes displayed complete discrimination of these hairpins (Fig. 20d).

To determine if dsDNA-regions with significant self-complementarity are particularly challenging targets for single-stranded  $\gamma$ PNA probes, we set out to study the dsDNA-invasion properties of  $^{Ser}\gamma$ PNA and Invader probes targeting DNA hairpins with low levels of self-complementarity.<sup>36</sup> Invader probes resulted in efficient, highly specific, but comparatively slow invasion of DNA hairpin targets, featuring complementary 14–16 base-paired stems and varying GC-content levels. Invasion was equally efficient but faster with the single-stranded  $^{Ser}\gamma$ PNA probes. Thus, three of six  $^{Ser}\gamma$ PNA probes and two of three Invader probes recognized the corresponding DNA hairpin targets with  $C_{50}$  values between 0.12–0.35  $\mu$ M (2.5 h

incubation, 37 °C) and 0.16–0.21  $\mu$ M (15 h incubation, 37 °C), respectively. Curiously, the denaturation profiles of three of six  $^{Ser}\gamma$ PNA probes in absence of cDNA were consistent with the formation of stable secondary structures in these sequence contexts. Additionally, at least two  $^{Ser}\gamma$ PNA probes displayed signs of aggregation and low solubility. Furthermore, DNA hairpins differing in sequence at one of 14–16 positions were poorly discriminated by  $^{Ser}\gamma$ PNA probes, whereas the corresponding Invader probes resulted in complete discrimination of these hairpins (Fig. 20e), once again underlining the advantageous impact of the stringency clamping effect that double-stranded probes exhibit. The lower binding specificity of the  $^{Ser}\gamma$ PNA probes also manifested itself in less specific labelling of corresponding chromosomal DNA regions (*vide infra*).

**Key point:** Optimized Invader probes allow for equally efficient recognition of mixed-sequence dsDNA targets with high levels of self-complementarity as corresponding INA probes, whereas single-stranded  $^{MP}\gamma$ PNA and fully modified LNA probes result in less efficient and less specific, but faster recognition. Invader probes result in equally efficient, substantially more specific, but slower recognition than single-stranded  $^{Ser}\gamma$ PNA probes when targeting mixed-sequence dsDNA regions with low levels of self-complementarity.

## Applications of Invader probes

Having demonstrated the effectiveness of second-generation Invaders probes for recognition of model mixed-sequence dsDNA targets using fluorescence and PAGE-based assays, we set out to establish proof-of-concept for Invader-mediated recognition of biological dsDNA targets.

In a first example,<sup>94</sup> we developed a 96-well plate sandwich assay based on Invader capture and signalling probes designed to recognize mixed-sequence dsDNA regions specific to three foodborne pathogens, *i.e.*, *Salmonella enterica*, *Campylobacter jejuni*, and *Escherichia coli*. Amine-terminated Invader capture probes were linked to the surfaces of 96-well plates and co-incubated with biotinylated Invader signalling probes and pathogen-specific dsDNA. Successful Invader-mediated recognition of the corresponding dsDNA targets was expected to result in the formation of two ternary complexes, *i.e.*, one in solution and one bound to the well (Fig. 21). Bound recognition complexes were detected using horseradish peroxidase (HRP)-conjugated streptavidin and subsequent incubation with an HRP substrate that forms a fluorescent product for convenient signal readout. Three sets of 14-mer Invader signalling and capture probes – each featuring three X-based hotspots – were designed to span the 28-mer pathogen-specific mixed-sequence dsDNA targets with GC-contents between 32% and 50%. As expected, the Invader capture and signalling probes were found to be relatively labile ( $\Delta T_m$  values between  $-5.0$  °C and  $+4.0$  °C, compared to the corresponding unmodified DNA duplexes), while individual probe strands displayed increased cDNA affinity ( $\Delta T_m$  values between  $+9.0$  °C and  $+19.0$  °C) resulting in favorable activation for dsDNA-recognition



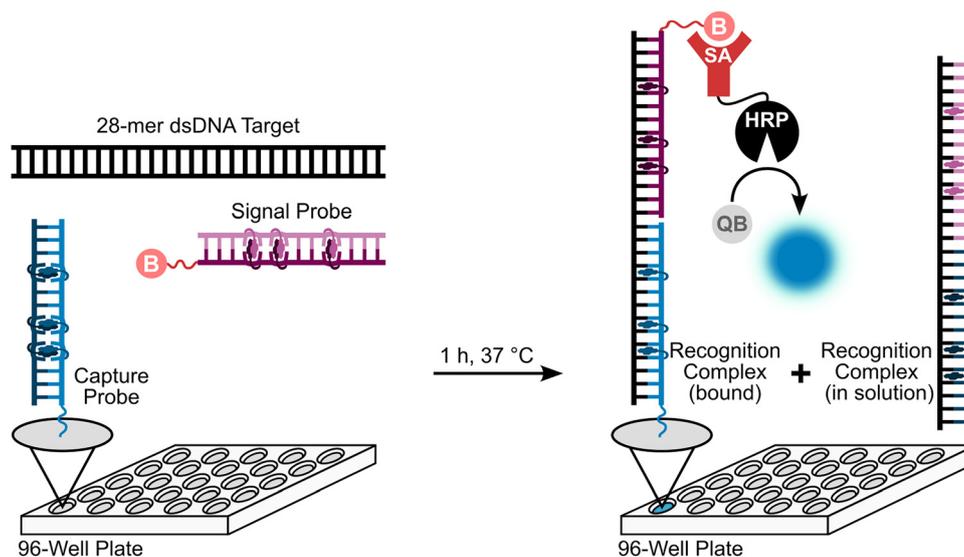


Fig. 21 Sandwich-based assay for detection of dsDNA targets specific to food pathogens.

(TA values between 21.0 °C and 40.0 °C). Incubation of Invader capture and signalling probes with the corresponding dsDNA targets (1 h, 37 °C) resulted in dose-dependent signal formation with a target sensitivity of 20–55 pM. In contrast, very weak signals were observed when signalling/capture probes designed for one species (*e.g.*, *E. coli*) were incubated with dsDNA targets specific to another species (*e.g.*, *Salmonella*), further

underscoring the high specificity of Invader-mediated dsDNA-recognition.

In a series of subsequent proof-of-principle studies,<sup>89–91</sup> we set out to utilize X-based Invader probes for recognition of chromosomal DNA in non-denaturing fluorescence *in situ* hybridization (nd-FISH) experiments (Fig. 22a). Unlike conventional FISH assays, which require chemical and/or heat-induced

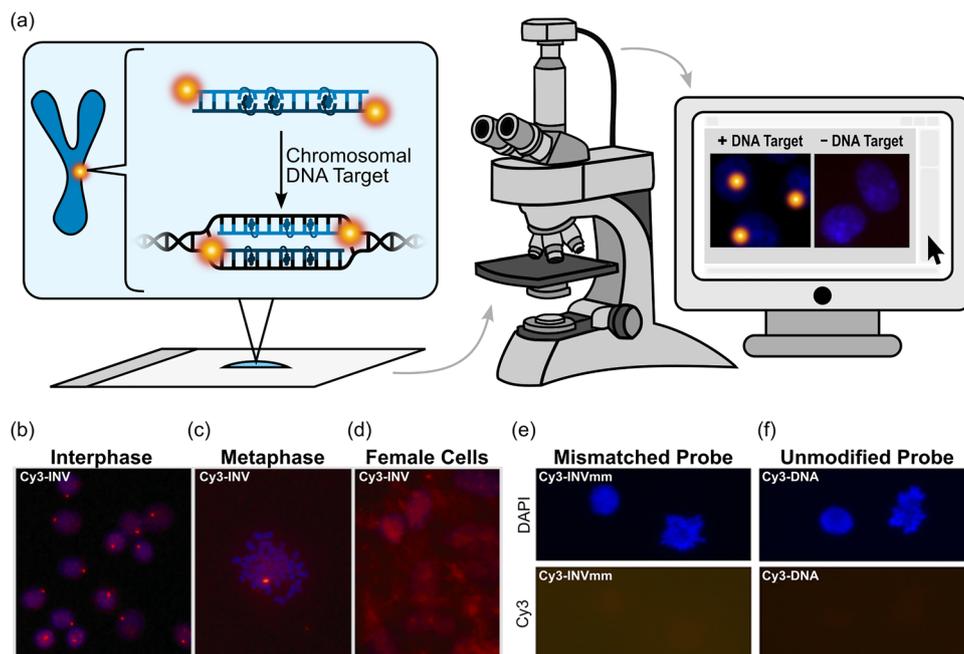


Fig. 22 Chromosomal detection in nd-FISH assays. (a) Illustration of assay. Representative fluorescence microscopy images from nd-FISH experiments using a target-specific Invader probe (Cy3-INV, 5'-Cy3-AGCCUGTGCCCTG:3'-TCGGACACGGGAC-Cy3) incubated with fixed nuclei from a male bovine cell line captured in (b) interphase and (c) metaphase, or (d) fixed nuclei from a female bovine cell line, or using (e) non-targeting Invader probe (Cy3-INVmm – sequence differing at three positions relative to target, 5'-Cy3-AGCGCUGAGGCCCTG:3'-TCGCGACTCCGGAC-Cy3), or (f) an unmodified DNA duplex with the same sequence as Cy3-INV (Cy3-DNA) incubated with fixed nuclei from the male bovine cell line under nd-FISH conditions. Images in panels (b)–(f) were reproduced from ref. 36, 89, and 91 from the Royal Society of Chemistry and MDPI.



denaturation of chromosomal DNA,<sup>95</sup> nd-FISH assays strive to map chromosomal loci under mild conditions. Towards this end, we selected a unique site within the highly repeated DYZ-1 satellite region ( $\sim 6 \times 10^4$  repeats) on the bovine (*Bos taurus*) Y chromosome as a model target. An initial set of four Cy3-labeled Invader probes – varying in length (11–14 base pairs) and number (2–3) and location of hotspots – were designed against the same target region in this site. As expected, the Invader probes were relatively labile ( $\Delta T_m$  values between  $-3.5$  °C and  $+1.0$  °C relative to the corresponding unmodified DNA duplexes), while the individual strands formed stable duplexes with cDNA ( $\Delta T_m$  values between  $+8.5$  °C and  $+15.0$  °C), resulting in a prominent driving force for dsDNA-recognition (TA values between  $+23.0$  °C and  $+26.5$  °C). The Cy3-labeled Invader probes were incubated with fixed interphase nuclei from a male bovine kidney cell line under non-denaturing conditions (3 h, 38.5 °C, 10 mM Tris-Cl, pH 8.0 and 1 mM EDTA), producing a single, punctate fluorescent signal that localized to the heterochromatic region, consistent with the expected localization of the DYZ-1 satellite target (Fig. 22b). One or two localized signals per nucleus – consistent with post- and premitotic nuclei – were observed when the probes were incubated with nuclei captured in metaphase (Fig. 22c). Each probe resulted in robust signals, with little variation between the probes. The high labeling coverage ( $\sim 90\%$ ), *i.e.*, the proportion of nuclei with localized signals, is noteworthy considering the high GC content of the target region ( $\sim 71\%$  GC).

Pre-treatment of nuclei with RNase A or proteinase K prior to incubation with the Invader probes, resulted in similar signals as in nuclei without pretreatment, while nuclei pre-treated with DNase I were devoid of localized signals, pointing to DNA as the target of the Invader probes.<sup>89</sup> Additionally, an absence of localized signals were observed when (i) fixed nuclei from a female bovine fibroblast cell line were incubated with Y-chromosome-specific Invader probes (Fig. 22d), or (ii) when Invader probes differing in sequence at three positions *vis-à-vis* the DYZ-1 target site were incubated with fixed nuclei from the male bovine kidney cell line (Fig. 22e). A Cy3-labeled but otherwise unmodified control of the Invader probes (Cy3-DNA) also failed to produce a signal (Fig. 22f). In concert, these results demonstrate that Invader-mediated recognition of mixed-sequence chromosomal DNA targets proceeds efficiently and in a highly specific manner.

In a subsequent study,<sup>91</sup> ten Invader probes based on 2'-O-(pyren-1-yl)methyl-RNA A/C/U monomers – varying in length (14–16 base pairs), hotspot content (20–30%), and GC-content (30–70%) – were designed to target a broader range of complementary sequences within the DYZ-1 region (Table 4). As discussed earlier, each of these Invader probes was found to be prominently activated for dsDNA-recognition, with all but two resulting in moderate to efficient recognition of the corresponding DNA hairpin targets ( $C_{50}$  values between 0.2  $\mu\text{M}$  and 4.1  $\mu\text{M}$ , Table 4). These Invader probes were incubated with fixed interphase nuclei from the male bovine kidney cell line under denaturing (d) or nd-FISH conditions. The d-FISH assay was expected to yield information about the maximal

recognition capacity of each probe, as kinetic barriers to accessing the chromosomal DNA target regions are minimized by the high incubation temperatures (5 min, 80 °C). Conversely, the nd-FISH experiments were designed to reveal if a probe can recognize the corresponding chromosomal DNA target at more physiologically relevant conditions. Three probes were found to recognize their DNA targets with excellent efficiency in the d-FISH assays (*i.e.*,  $\sim 90\%$  of the analyzed interphase nuclei displaying a single, intense, punctate signal against a low level of background), five probes displayed single punctate signals in 40–60% of the analyzed nuclei, while two probes failed to yield acceptable signal profiles (Table 4). The Invader probes largely retained their signaling characteristics under nd-FISH conditions (Table 4). Thus, three probes yielded single, intense, and punctate signals against a low background in 85–90% of the analyzed nuclei (Fig. 23), four probes displayed moderately intense signals in 20–30% of the nuclei, while three probes did not produce discernable signals (Table 4). In other words, most of the studied Invader probes resulted in acceptable-to-excellent recognition of chromosomal DNA targets under d-FISH and nd-FISH conditions.

A Spearman's rank-order correlation analysis revealed that signaling performance in d-FISH and nd-FISH assays correlated with the hotspot density of the probes and  $C_{50}$  values from the DHP assays (Fig. 15). Moreover, correlations approaching statistical significance were observed between nd-FISH (but not d-FISH) signaling performance and TA and  $\Delta G_{\text{rec}}^{310}$  values, indicating that these metrics have predictive value for nd-FISH, but not d-FISH, assay performance.

Having identified hotspot density as the key parameter for successful dsDNA-recognition, the hotspot density in three Invader probes displaying poor-to-moderate signaling characteristics under nd-FISH conditions (**INV6**, **INV8** and **INV9**) was increased from  $\sim 20\%$  to  $\sim 30\%$  (**OPT6**, **OPT8**, **OPT9**) (Table 4). As predicted, the optimized probes displayed more prominent driving forces for recognition of dsDNA targets than the parent probes (TA values between 43.0–55.5 °C *vs.* 13.5–26.5 °C, respectively;  $\Delta G_{\text{rec}}^{310}$  values between  $-93$  kJ mol<sup>-1</sup> and  $-59$  kJ mol<sup>-1</sup> *vs.* between  $-52$  kJ mol<sup>-1</sup> and  $-19$  kJ mol<sup>-1</sup>, respectively) and lower  $C_{50}$  values, whilst still displaying excellent binding specificity. In agreement with these trends, improved signaling characteristics, relative to the parent probes, were observed for the optimized probes. Thus, 75–90% of the nuclei displayed prominent, single, and punctate signals when optimized probes were used under d-FISH conditions compared to  $\sim 60\%$  of nuclei when parent probes were used (Table 4). Along similar lines,  $\sim 85\%$ ,  $\sim 75\%$  and  $\sim 25\%$  of the nuclei displayed high-quality signals when the three optimized probes were used under nd-FISH conditions, respectively, as compared to 0–25% with the parent probes (Table 4). This indicates that increasing the hotspot density is an effective strategy to improve the signalling characteristics of Invader probes.

The nd-FISH signalling characteristics of six TAMRA-labelled <sup>ser</sup> $\gamma$ PNA probes, targeting the sense or antisense strand of the same regions as the three most promising Cy3-labelled Invader probes, were determined next.<sup>36</sup> Two of the <sup>ser</sup> $\gamma$ PNA probes resulted in the formation of single punctate signals in



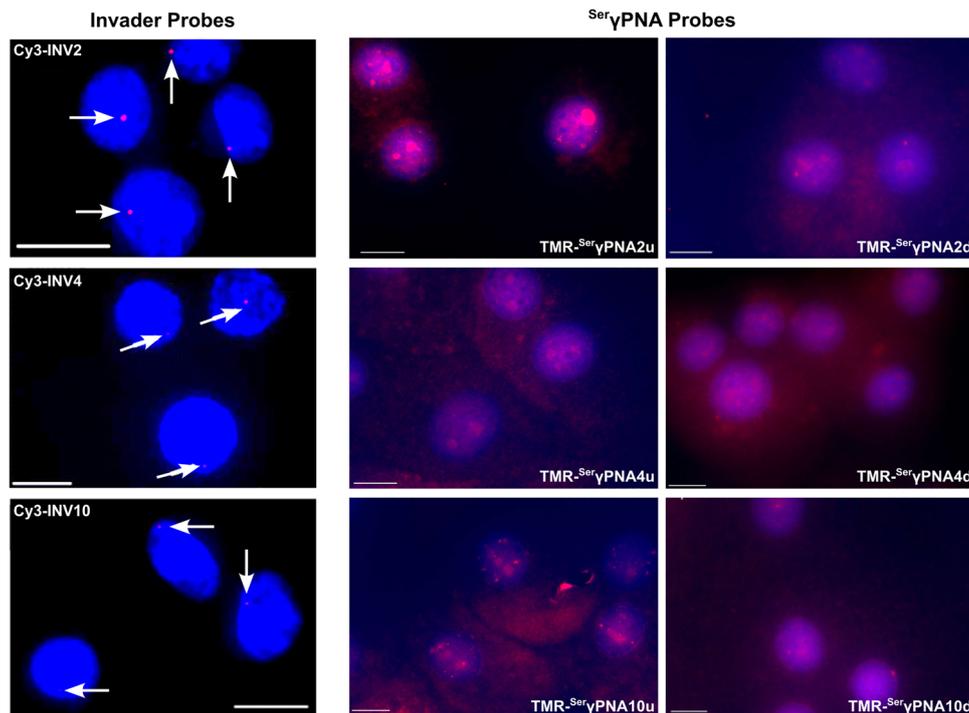


Fig. 23 Representative fluorescence microscopy images from nd-FISH experiments in which target-specific Invader probes (Cy3-INV2, Cy3-INV4, and Cy3-INV10) and the corresponding  $\text{Ser}\gamma\text{PNA}$  probes (TMR- $\text{Ser}\gamma\text{PNA}2\text{u}$ , TMR- $\text{Ser}\gamma\text{PNA}2\text{d}$ , TMR- $\text{Ser}\gamma\text{PNA}4\text{u}$ , TMR- $\text{Ser}\gamma\text{PNA}4\text{d}$ , TMR- $\text{Ser}\gamma\text{PNA}10\text{u}$ , and TMR- $\text{Ser}\gamma\text{PNA}10\text{d}$ ) were incubated with fixed nucleic acid from a male bovine kidney cell line. Reproduced from ref. 36 with permission from the Royal Society of Chemistry.

~70% of interphase nuclei from the male bovine kidney cell line but with a stronger non-specific background than with the Invader probes, which produced single punctate signals in >85% of the interphase nuclei (Fig. 23). The remaining  $\text{Ser}\gamma\text{PNA}$  probes resulted in no or minimal target-specific recognition; instead, multiple signals and/or intense diffuse background were observed. Thus, despite the two probe chemistries displaying comparable binding affinities against the DNA hairpin targets (see the section “Comparison of second-generation invader probes with other DNA-targeting probes”), considerable differences were observed in the nd-FISH experiments.

Invader probes have also been evaluated for their ability to target telomeric DNA.<sup>92</sup> Telomeres are guanosine-rich tandem

repeats (5'-TTAGGG-3', in vertebrates) capping chromosome termini and have central roles in aging, age-associated diseases, and cancer (Fig. 24).<sup>96,97</sup> Accordingly, there has been considerable interest in developing fluorescence-based diagnostic tools that enable detection of telomeres.<sup>7,98–102</sup> Towards this end, a 16-mer Cy3-labeled Invader probe with three X-based hotspots was designed. The probe displayed the expected thermal denaturation properties, *i.e.*, the probe duplex was less stable than the corresponding duplexes between individual probe strands and cDNA, although the stability differences were less pronounced (TA = 19.5 °C). Nonetheless, incubation of this Invader probe with isolated nuclei from a male bovine kidney cell line, produced pairs of bright foci on the termini of

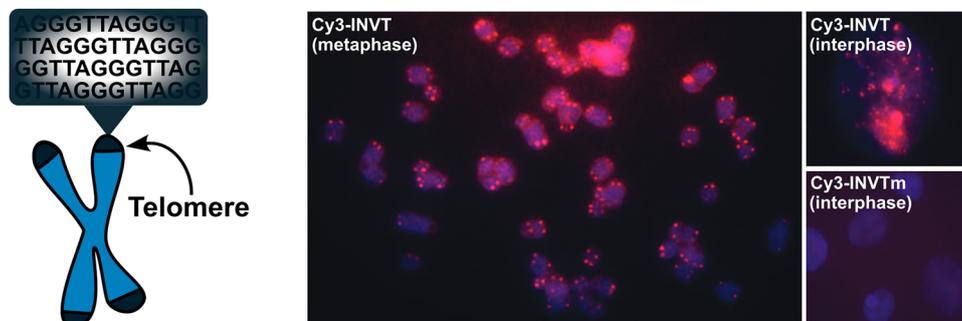


Fig. 24 Illustration of telomere target and representative fluorescence microscopy images from nd-FISH experiments in which target-specific Invader probe (Cy3-INVt) and single-mismatched Invader probe (Cy3-INVtm) were incubated with isolated nuclei in metaphase or interphase spreads. Reproduced from ref. 92 with permission from the Royal Society of Chemistry.



most metaphasic chromosomes in nd-FISH experiments (Fig. 24). In contrast, a corresponding 16-mer Invader probe differing in sequence at one position relative to the telomeric repeat, did not result in signal formation (Fig. 24).<sup>92</sup>

**Key point:** Appropriately designed Invader probes have found utility in diagnostic assays and enable highly specific and sequence-unrestricted recognition of chromosomal DNA targets at denaturing and non-denaturing conditions.

## Conclusion and future outlook

What began as a curiosity-driven study with no clear application beyond the pursuit of improved LNA analogues (*i.e.*, synthesis and biophysical characterization of ONs modified with N2'-functionalized 2'-amino- $\alpha$ -L-LNA monomers), led to an unexpected discovery: DNA duplexes with +1 interstrand zippers of intercalator-functionalized nucleotides are labile, while individual strands display remarkable affinity for complementary DNA. This serendipitous finding evolved into a multi-year effort that yielded a new and distinctive class of strand-invading probes capable of sequence-unrestricted recognition of dsDNA target regions – the Invader probes. Much of our work has focused on the chemical refinement of Invader probes through optimization of the monomers<sup>76–78,82,103</sup> and probe architectures,<sup>35–38,104–110</sup> which will be the subject of a separate account. In the process, more than ten graduate students and fifteen undergraduate students have received training in synthetic organic chemistry, nucleic acid chemistry, and molecular biology – most now contributing their expertise in the pharmaceutical industry.

Although key evidence already supports the utility of Invader probes – most notably their ability to target specific chromosomal DNA regions under nd-FISH conditions – much remains to be explored. Demonstrating efficient nuclear uptake and gene knockdown in cell culture would provide important proof to establish Invader probes as a solution to the long-standing challenge of robust and sequence-unrestricted dsDNA-recognition. Unpublished studies – using sub-optimal, sparsely modified, first-generation Invader probes – have indicated that inhibition of *in vitro* transcription and gene knockdown in luciferase reporter assays in cell cultures is possible, suggesting that the concept is viable and primed for further exploration using optimally designed, next-generation Invader probes in collaboration with domain experts. Other applications of Invader probes seem possible including their use as PCR blockers for enrichment of rare mutants, artificial restriction enzymes, and non-protein-based gene editing constructs.

In summary, Invader probes reveal that directed intercalation can be harnessed to overcome the energetic penalty of duplex invasion to achieve sequence-unrestricted access to double-stranded DNA without reliance on proteins.

## Conflicts of interest

PJH is an inventor on patents pertaining to Invader probes, which have been issued to the University Idaho.

## Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

## Acknowledgements

PJH is indebted to many former and current students, colleagues, and collaborators who have contributed to this project over the years, most of whom are mentioned in the referenced works. PJH also appreciates financial support for this work over the years from multiple funding agencies including award GM-088697 from the National Institute of General Medical Sciences, National Institutes of Health; awards IF13-001, IF14-012 and IGEM-004 from the Higher Education Research Council, Idaho State Board of Education; the BANTech Center at the University of Idaho; Idaho NSF EPSCoR and an Institutional Development Award (IDEA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant #P20GM103408. The authors also wish to thank the reviewers of this manuscript for many helpful comments.

## References

- 1 S. Naeem, J. Zhang, Y. Zhang and Y. Wang, *Mol. Ther. – Nucleic Acids*, 2025, **36**, 102440.
- 2 P. B. Dervan and B. S. Edelson, *Curr. Opin. Struct. Biol.*, 2003, **13**, 284–299.
- 3 Y. Kawamoto, T. Bando and H. Sugiyama, *Bioorg. Med. Chem.*, 2018, **26**, 1393–1411.
- 4 M. Duca, P. Vekhoff, K. Oussedik, L. Halby and P. B. Arimondo, *Nucleic Acids Res.*, 2008, **36**, 5123–5138.
- 5 Y. Mikame and A. Yamayoshi, *Pharmaceutics*, 2023, **15**, 2515.
- 6 N. Brodyagin, M. Katkevics, V. Kotikam, C. A. Ryan and E. Rozners, *Beilstein J. Org. Chem.*, 2021, **17**, 1641–1688.
- 7 Y. Kawamoto, A. Sasaki, A. Chandran, K. Hashiya, S. Ide, T. Bando, K. Maeshima and H. Sugiyama, *J. Am. Chem. Soc.*, 2016, **138**, 14100–14107.
- 8 T. Bentin, H. J. Larsen and P. E. Nielsen, *Biochemistry*, 2003, **42**, 13987–13995.
- 9 K. Kaihatsu, R. H. Shah, X. Zhao and D. R. Corey, *Biochemistry*, 2003, **42**, 13996–14003.
- 10 Y. Hari, S. Obika and T. Imanishi, *Eur. J. Org. Chem.*, 2012, 2875–2887.
- 11 V. V. Filichev, M. C. Nielsen, N. Bomholt, C. H. Jessen and E. B. Pedersen, *Angew. Chem.*, 2006, **118**, 5437–5441.
- 12 L. De Napoli, A. Messere, D. Montesarchio, A. Pepe, G. Piccialli and M. Varra, *J. Org. Chem.*, 1997, **62**, 9024–9030.
- 13 D. A. Rusling, V. E. C. Powers, R. T. Ranasinghe, Y. Wang, S. D. Osborne, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2005, **33**, 3025–3032.
- 14 A. E. Modell, D. Lim, T. M. Nguyen, V. Sreekanth and A. Choudhary, *Trends Pharmacol. Sci.*, 2022, **43**, 151–161.



- 15 S. Sethi, H. Zumila, Y. Watanabe, J. Mo and K. Fujimoto, *Bioorg. Med. Chem. Lett.*, 2024, **98**, 129597.
- 16 S. A. Thadke, V. M. Hridya, J. D. R. Perera, R. R. Gil, A. Mukherjee and D. H. Ly, *Commun. Chem.*, 2018, **1**, 1–10.
- 17 H. Zheng, I. Botos, V. Clause, H. Nikolayevskiy, E. E. Rastede, M. F. Fouz, S. J. Mazur and D. H. Appella, *Nucleic Acids Res.*, 2021, **49**, 713–725.
- 18 S. Rapireddy, G. He, S. Roy, B. A. Armitage and D. H. Ly, *J. Am. Chem. Soc.*, 2007, **129**, 15596–15600.
- 19 A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil and D. H. Ly, *J. Am. Chem. Soc.*, 2006, **128**, 10258–10267.
- 20 R. Bahal, B. Sahu, S. Rapireddy, C.-M. Lee and D. H. Ly, *ChemBioChem*, 2012, **13**, 56–60.
- 21 P. R. Bohländer, T. Vilaivan and H.-A. Wagenknecht, *Org. Biomol. Chem.*, 2015, **13**, 9223–9230.
- 22 H. Kaur, B. R. Babu and S. Maiti, *Chem. Rev.*, 2007, **107**, 4672–4697.
- 23 P. H. Hagedorn, R. Persson, E. D. Funder, N. Albæk, S. L. Diemer, D. J. Hansen, M. R. Møller, N. Papargyri, H. Christiansen, B. R. Hansen, H. F. Hansen, M. A. Jensen and T. Koch, *Drug Discovery Today*, 2018, **23**, 101–114.
- 24 K. M. L. Hertoghs, *Nucleic Acids Res.*, 2003, **31**, 5817–5830.
- 25 R. Beane, S. Gabillet, C. Montallier, K. Arar and D. R. Corey, *Biochemistry*, 2008, **47**, 13147–13149.
- 26 E. M. Zaghoul, A. S. Madsen, P. M. D. Moreno, I. I. Oprea, S. El-Andaloussi, B. Bestas, P. Gupta, E. B. Pedersen, K. E. Lundin, J. Wengel and C. I. E. Smith, *Nucleic Acids Res.*, 2011, **39**, 1142–1154.
- 27 O. Saher, E. M. Zaghoul, T. Umek, D. W. Hagey, N. Mozafari, M. B. Danielsen, A. S. Gouda, K. E. Lundin, P. T. Jørgensen, J. Wengel, C. I. E. Smith and R. Zain, *Nucleic Acid Ther.*, 2023, **33**, 117–131.
- 28 A. G. Olsen, O. Dahl, A. B. Petersen, J. Nielsen and P. E. Nielsen, *Artif. DNA PNA XNA*, 2011, **2**, 33–37.
- 29 I. V. Kutyavin, R. L. Rhinehart, E. A. Lukhtanov, V. V. Gorn, R. B. Meyer and H. B. Gamper, *Biochemistry*, 1996, **35**, 11170–11176.
- 30 J. Lohse, O. Dahl and P. E. Nielsen, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 11804–11808.
- 31 V. V. Demidov, E. Protozanova, K. I. Izvolsky, C. Price, P. E. Nielsen and M. D. Frank-Kamenetskii, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5953–5958.
- 32 T. Bryld, T. Højland and J. Wengel, *Chem. Commun.*, 2004, 1064–1065.
- 33 V. V. Filichev, U. B. Christensen, E. B. Pedersen, B. R. Babu and J. Wengel, *ChemBioChem*, 2004, **5**, 1673–1679.
- 34 H. Asanuma, R. Niwa, M. Akahane, K. Murayama, H. Kashida and Y. Kamiya, *Bioorg. Med. Chem.*, 2016, **24**, 4129–4137.
- 35 R. G. Emehiser and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2020, **18**, 1359–1368.
- 36 R. G. Emehiser, K. Dhuri, C. Shepard, S. Karmakar, R. Bahal and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2022, **20**, 8714–8724.
- 37 M. E. Everly, R. G. Emehiser and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2025, **23**, 619–628.
- 38 M. E. Everly and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2026, **24**, 319–323.
- 39 P. Wittung, S. K. Kim, O. Buchardt, P. Nielsen and B. Norden, *Nucleic Acids Res.*, 1994, **22**, 5371–5377.
- 40 V. Marin, H. F. Hansen, T. Koch and B. A. Armitage, *J. Biomol. Struct. Dyn.*, 2004, **21**, 841–850.
- 41 M. Hibino, Y. Aiba and O. Shoji, *Chem. Commun.*, 2020, **56**, 2546–2549.
- 42 A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen and J. Wengel, *Tetrahedron*, 1998, **54**, 3607–3630.
- 43 V. K. Rajwanshi, A. E. Håkansson, M. D. Sørensen, S. Pitsch, S. K. Singh, R. Kumar, P. Nielsen and J. Wengel, *Angew. Chem., Int. Ed.*, 2000, **39**, 1656–1659.
- 44 J. Wengel, *Org. Biomol. Chem.*, 2004, **2**, 277–280.
- 45 P. J. Hrdlicka, B. R. Babu, M. D. Sørensen and J. Wengel, *Chem. Commun.*, 2004, 1478–1479.
- 46 P. J. Hrdlicka, B. R. Babu, M. D. Sørensen, N. Harrit and J. Wengel, *J. Am. Chem. Soc.*, 2005, **127**, 13293–13299.
- 47 T. S. Kumar, J. Wengel and P. J. Hrdlicka, *ChemBioChem*, 2007, **8**, 1122–1125.
- 48 T. S. Kumar, A. S. Madsen, M. E. Østergaard, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2008, **73**, 7060–7066.
- 49 M. E. Østergaard and P. J. Hrdlicka, *Chem. Soc. Rev.*, 2011, **40**, 5771–5788.
- 50 N. E. Poopeiko, M. Juhl, B. Vester, M. D. Sørensen and J. Wengel, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 2285–2290.
- 51 T. S. Kumar, A. S. Madsen, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2006, **71**, 4188–4201.
- 52 T. S. Kumar, A. S. Madsen, M. E. Østergaard, S. P. Sau, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2009, **74**, 1070–1081.
- 53 N. K. Andersen, B. A. Anderson, J. Wengel and P. J. Hrdlicka, *J. Org. Chem.*, 2013, **78**, 12690–12702.
- 54 P. J. Hrdlicka, T. S. Kumar and J. Wengel, *Chem. Commun.*, 2005, 4279–4281.
- 55 U. B. Christensen and E. B. Pedersen, *Nucleic Acids Res.*, 2002, **30**, 4918–4925.
- 56 V. V. Filichev, K. M. H. Hilmy, U. B. Christensen and E. B. Pedersen, *Tetrahedron Lett.*, 2004, **45**, 4907–4910.
- 57 K. Yamana, R. Iwase, S. Furutani, H. Tsuchida, H. Zako, T. Yamaoka and A. Murakami, *Nucleic Acids Res.*, 1999, **27**, 2387–2392.
- 58 N. Kalra, B. R. Babu, V. S. Parmar and J. Wengel, *Org. Biomol. Chem.*, 2004, **2**, 2885–2887.
- 59 V. A. Korshun, D. A. Stetsenko and M. J. Gait, *J. Chem. Soc., Perkin Trans. 1*, 2002, 1092–1104.
- 60 C. Dohno and I. Saito, *ChemBioChem*, 2005, **6**, 1075–1081.
- 61 H. Asanuma, T. Fujii, T. Kato and H. Kashida, *J. Photochem. Photobiol., C*, 2012, **13**, 124–135.
- 62 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**, 9560–9570.
- 63 S. P. Sau, T. S. Kumar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2010, **8**, 2028.



- 64 F. M. Winnik, *Chem. Rev.*, 1993, **93**, 587–614.
- 65 C. B. Nielsen, M. Petersen, E. B. Pedersen, P. E. Hansen and U. B. Christensen, *Bioconjugate Chem.*, 2004, **15**, 260–269.
- 66 H. Ihmels and D. Otto, *Supramolecular Dye Chemistry*, Springer, Berlin, Heidelberg, 2005, pp. 161–204.
- 67 Ö. Persil and N. V. Hud, *Trends Biotechnol.*, 2007, **25**, 433–436.
- 68 S. C. Jain, C. Tsai and H. M. Sobell, *J. Mol. Biol.*, 1977, **114**, 317–331.
- 69 L. D. Williams, M. Egli, Q. Gao and A. Rich, *Structure and Function: Nucleic Acids*, Adenine Press, Albany, NY, USA, 1992, vol. 1, pp. 107–125.
- 70 L. P. Reynaldo, A. V. Vologodskii, B. P. Neri and V. I. Lyamichev, *J. Mol. Biol.*, 2000, **297**, 511–520.
- 71 V. V. Filichev, B. Vester, L. H. Hansen and E. B. Pedersen, *Nucleic Acids Res.*, 2005, **33**, 7129–7137.
- 72 M. Nakamura, Y. Fukunaga, K. Sasa, Y. Ohtoshi, K. Kanaori, H. Hayashi, H. Nakano and K. Yamana, *Nucleic Acids Res.*, 2005, **33**, 5887–5895.
- 73 K. Yamana, Y. Ohashi, K. Nunota, M. Kitamura, H. Nakano, O. Sangen and T. Shimidzu, *Tetrahedron Lett.*, 1991, **32**, 6347–6350.
- 74 S. Karmakar, B. A. Anderson, R. L. Rathje, S. Andersen, T. B. Jensen, P. Nielsen and P. J. Hrdlicka, *J. Org. Chem.*, 2011, **76**, 7119–7131.
- 75 B. S. Ross, R. H. Springer, Z. Tortorici and S. Dimock, *Nucleosides Nucleotides*, 1997, **16**, 1641–1643.
- 76 S. Karmakar, A. S. Madsen, D. C. Guenther, B. C. Gibbons and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2014, **12**, 7758–7773.
- 77 S. Karmakar, D. C. Guenther, B. C. Gibbons and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2017, **15**, 9362–9371.
- 78 S. Karmakar, T. Horrocks, B. C. Gibbons, D. C. Guenther, R. Emehiser and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2019, **17**, 609–621.
- 79 S. Karmakar, D. C. Guenther and P. J. Hrdlicka, *J. Org. Chem.*, 2013, **78**, 12040–12048.
- 80 M. Nakamura, Y. Shimomura, Y. Ohtoshi, K. Sasa, H. Hayashi, H. Nakano and K. Yamana, *Org. Biomol. Chem.*, 2007, **5**, 1945–1951.
- 81 S. Al-Rawi, C. Ahlborn and C. Richert, *Org. Lett.*, 2005, **7**, 1569–1572.
- 82 B. A. Anderson, J. J. Onley and P. J. Hrdlicka, *J. Org. Chem.*, 2015, **80**, 5395–5406.
- 83 D. P. C. McGee, A. Vaughn-Settle, C. Vargeese and Y. Zhai, *J. Org. Chem.*, 1996, **61**, 781–785.
- 84 M. Manoharan, K. Tivel, M. Zhao, K. Nafisi and N. Thomas, *J. Phys. Chem.*, 1995, **99**, 17461–17472.
- 85 J. N. Wilson, Y. Cho, S. Tan, A. Cuppoletti and E. T. Kool, *ChemBioChem*, 2008, **9**, 279–285.
- 86 N. N. Dioubankova, A. D. Malakhov, D. A. Stetsenko, M. J. Gait, P. E. Volynsky, R. G. Efremov and V. A. Korshun, *ChemBioChem*, 2003, **4**, 841–847.
- 87 I. V. Astakhova, A. D. Malakhov, I. A. Stepanova, A. V. Ustinov, S. L. Bondarev, A. S. Paramonov and V. A. Korshun, *Bioconjugate Chem.*, 2007, **18**, 1972–1980.
- 88 I. V. Astakhova, A. V. Ustinov, V. A. Korshun and J. Wengel, *Bioconjugate Chem.*, 2011, **22**, 533–539.
- 89 B. A. Didion, S. Karmakar, D. C. Guenther, S. P. Sau, J. P. Versteegen and P. J. Hrdlicka, *ChemBioChem*, 2013, **14**, 1534–1538.
- 90 D. C. Guenther, G. H. Anderson, S. Karmakar, B. A. Anderson, B. A. Didion, W. Guo, J. P. Versteegen and P. J. Hrdlicka, *Chem. Sci.*, 2015, **6**, 5006–5015.
- 91 C. P. Shepard, R. G. Emehiser, S. Karmakar and P. J. Hrdlicka, *Molecules*, 2023, **28**, 127.
- 92 R. G. Emehiser, E. Hall, D. C. Guenther, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2020, **18**, 56–65.
- 93 V. V. Demidov and M. D. Frank-Kamenetskii, *Trends Biochem. Sci.*, 2004, **29**, 62–71.
- 94 B. Denn, S. Karmakar, D. C. Guenther and P. J. Hrdlicka, *Chem. Commun.*, 2013, **49**, 9851–9853.
- 95 J. M. Levsky and R. H. Singer, *J. Cell Sci.*, 2003, **116**, 2833–2838.
- 96 J. W. Shay, *Cancer Discovery*, 2016, **6**, 584–593.
- 97 P. Martínez and M. A. Blasco, *J. Cell Biol.*, 2017, **216**, 875–887.
- 98 H. H. Pham, C. T. Murphy, G. Sureshkumar, D. H. Ly, P. L. Opresko and B. A. Armitage, *Org. Biomol. Chem.*, 2014, **12**, 7345–7354.
- 99 A. Orenstein, A. S. Berlyoung, E. E. Rastede, H. H. Pham, E. Fouquerel, C. T. Murphy, B. J. Leibowitz, J. Yu, T. Srivastava, B. A. Armitage and P. L. Opresko, *Molecules*, 2017, **22**, 2117.
- 100 H. Ma, A. Naseri, P. Reyes-Gutierrez, S. A. Wolfe, S. Zhang and T. Pederson, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 3002–3007.
- 101 P. M. Lansdorp, *J. Cell Biol.*, 1997, **139**, 309–312.
- 102 A. S. Boutorine, D. S. Novopashina, O. A. Krasheninina, K. Nozeret and A. G. Venyaminova, *Molecules*, 2013, **18**, 15357–15397.
- 103 B. A. Anderson and P. J. Hrdlicka, *Bioorg. Med. Chem. Lett.*, 2015, **25**, 3999–4004.
- 104 B. A. Anderson, S. Karmakar and P. J. Hrdlicka, *Molecules*, 2015, **20**, 13780–13793.
- 105 D. C. Guenther, S. Karmakar and P. J. Hrdlicka, *Chem. Commun.*, 2015, **51**, 15051–15054.
- 106 B. A. Anderson and P. J. Hrdlicka, *J. Org. Chem.*, 2016, **81**, 3335–3346.
- 107 S. P. Adhikari, R. G. Emehiser, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2019, **17**, 8795–8799.
- 108 D. C. Guenther, R. G. Emehiser, A. Inskeep, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2020, **18**, 4645–4655.
- 109 S. P. Adhikari, P. Vukelich, D. C. Guenther, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2021, **19**, 9276–9290.
- 110 S. P. Adhikari, S. Karmakar and P. J. Hrdlicka, *Org. Biomol. Chem.*, 2022, **20**, 1019–1030.

