HHS Public Access

Author manuscript

Chembiochem. Author manuscript; available in PMC 2023 May 24.

Published in final edited form as:

Chembiochem. 2012 January 02; 13(1): 56–60. doi:10.1002/cbic.201100646.

Sequence-Unrestricted, Watson-Crick Recognition of Double Helical B-DNA by (*R*)-MiniPEG-γPNAs

Dr. Raman Bahal [Prof.],

Dr. Bichismita Sahu,

Srinivas Rapireddy,

Chong-Min Lee,

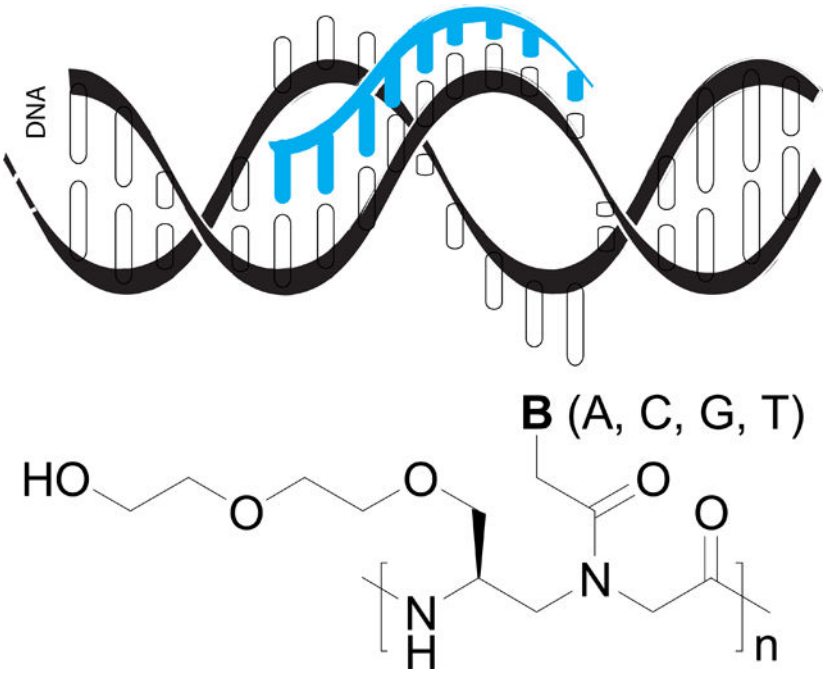
Danith H. Ly* **[Prof.]**

Department of Chemistry and Center for Nucleic Acids, Science and Technology (CNAST), Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213 (USA)

Entry for the Table of Contents

Chiral γPNAs containing miniPEG side-chains can invade any sequence of double helical B-form DNA, with the recognition occurring through direct Watson-Crick base-pairing.

Graphical Abstract



*Fax: (+)1 412-268-5579, dly@andrew.cmu.edu.

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

Keywords

DNA recognition; γPNA; strand invasion; Watson-Crick base-pairing

Development of general principles for designing molecules to bind sequence-specifically to double-stranded DNA (dsDNA) has been a long-sought goal of bioorganic chemistry and molecular biology.[1–3] Pursuit of this goal, in the past, has generally been focused on the minor and major grooves—in large part, because of the ease of accessibility of the chemical groups that reside on these external parts of the double helix and the precedence for their recognition in nature.[2] It was long recognized that while Watson-Crick (W-C) base-pairing provides a more direct and specific means for establishing sequence-specific interactions with nucleic acid biopolymers such as DNA and RNA, it would be difficult to do so with intact double helical DNA because of the preexisting base-pairs.[4] This effort has so far led to the development of several major classes of antigene molecules, with the like of triplex-forming oligonucleotides,[5–7] minor-groove binding polyamides,[8–11] and major-groove binding zinc-finger peptides.[12–16] While they can be designed to bind sequence-specifically to dsDNA, there are still remaining issues with sequence selection, specificity and/or target length that have not yet been completely resolved[8, 13, 17–19]— although some progress has been made in recent years.

Over the past two decades, peptide nucleic acids (PNAs)[20]—a particular class of nucleic acid mimics comprised of a pseudopeptide backbone (Scheme 1A)—have been shown to be capable of invading dsDNA.[21] This finding is significant because, contrary to the traditional belief, it demonstrates that DNA double helix is relatively dynamic at physiological temperatures and that W-C base-pairing interactions can be established with intact dsDNA. Though promising as antigene reagents, because of the specificity of recognition and generality in sequence design, PNA binding is presently limited to mostly homopurine[22] and homopyrimidine targets.[23] Mixed-sequence PNAs have been shown to be capable of invading topologically constrained supercoiled plasmid DNA,[24–27] conformationally perturbed regions of genomics DNA[28] and duplex termini;[29] however, they are unable to invade the interior regions of double helical B-form DNA (B-DNA)—the most stable form of DNA double helix. “Tail-clamp”[30, 31] and “double-duplex invasion”[32] strategies have subsequently been developed, enabling mixed-sequence PNAs to invade B-DNA, but they are not without limitations.[33]

Recently we showed that mixed-sequence PNAs, when preorganized into a right-handed helix by installing an (*S*)-Me stereogenic center at the γ-backbone (Scheme 1B), can invade B-DNA.[34] However, all of our studies so far are limited to a few selected sequences due to the poor water solubility and propensity of these first-generation $^{S\text{-Me}}$γPNAs to aggregate and adhere to surfaces and other macromolecules, including DNA, in a nonspecific manner. This problem is exacerbated with increasing G/C content, making it difficult to characterize their invasion properties over a wide concentration range and broad sequence space. To rectify this problem, we replaced the methyl group with diethylene glycol (or ‘MiniPEG’, MP [Scheme 1C]). MP was chosen because of its relatively small size and hydrophilic nature, and because it has been shown to be effective in imparting water solubility and

biocompatibility to non-biological systems.[35] Our initial study revealed that inclusion of MP indeed significantly improves the water solubility and biocompatibility of PNAs,[36] but whether these MP-modified γPNAs ($^{R\text{-MP}}$γPNAs) can invade B-DNA has not yet been determined. Herein we show that $^{R\text{-MP}}$γPNAs can invade any sequence of double helical B-DNA, ranging from 0 to 100% G/C content, with the recognition occurring in a highly sequence specificity through Watson-Crick base-pairing.

To assess binding, we performed gel-shift assays, comparing $^{R\text{-MP}}$γPNA1 to that of $^{S\text{-Me}}$γPNA1 and PNA1 (Table 1). A 291-bp DNA fragment containing an internal binding site was chosen as a target (Figure 1A, PM) because initial study showed that it is short enough to impart electrophoretic separation upon binding of γPNAs and long enough not to interfere with the invasion efficiency (Figure S1, Supporting Information). The target was incubated with different concentrations of oligos in 10 mM sodium phosphate (NaPi) buffer for 16 hrs, followed by electrophoretic separation and staining with SYBR-Gold for visualization. Our result showed that not only could $^{R\text{-MP}}$γPNA1 bind dsDNA, but it did so with greater efficiency than $^{S\text{-Me}}$γPNA, as evidenced by the relative intensity of the retarded bands (Figure 1B; compare lanes 5–7 to lanes 8–10, respectively). The dissociation constant (K_d) of $^{R\text{-MP}}$γPNA1 to **PM** was determined to be 3.7 (±0.2) x 10^{-7} M. However, under identical conditions, no binding was observed for the achiral PNA (lanes 2–4). This result is consistent with the earlier finding.[37] Even at a high (100:1) oligo-to-DNA strand ratio, at which point all DNA incubated with $^{S\text{-Me}}$γPNA (lane 7) and PNA1 (lane 4) had disappeared, presumably due to aggregation and nonspecific binding of $^{S\text{-Me}}$γPNA and PNA to DNA, the presence of large excess $^{R\text{-MP}}$γPNA1 had no effect on the mobility or intensity of the retarded band (lane 10). This result highlights the importance of MP at the γ-backbone in suppressing aggregation and nonspecific binding.

Formation of the complex, in this case, occurred in a highly sequence-specific manner. No incubation of DNA, whether containing a single-base mismatch in the middle (**MM1**) or towards one end (**MM2** or **MM3**), resulted in formation of the retarded band (Figure 1C). This finding was further corroborated by DNase-I footprinting, which revealed protection at the expected binding site (Figure 2A). The footprinting pattern was only observed with the perfect-match (lanes 2–4) and not with the single-base mismatches (lanes 5 and 6). The binding mode of $^{R\text{-MP}}$γPNA1 was confirmed by diethyl pyrocarbonate (DEPC)-chemical probing assay (Figure 2B),[38, 39] which revealed selective cleavage at the adenine and, to a smaller extent, guanine sites on the homologous DNA strand, directly across from the binding site, following piperidine treatment. This result is consistent with binding occurring through a strand invasion mechanism.

Next, we determined the effects of temperature and ionic strength on the invasion efficiency. Our result showed that the invasion efficiency is strongly dependent on temperature (Figure 3A). This was an expected finding, because temperature has a direct effect on the rate of base-pair breathing (or opening). Strand invasion was extremely inefficient at 4 °C, in which case less than 10% of the invasion complex was formed after a 24 hr-incubation. The rate of strand invasion dramatically increased at 37 °C, reaching equilibrium within ~ 2 hrs. The rate obeys peudo-first order kinetics, with $k_{ps,\ 37\ °C} = 0.025\ min^{-1}$ (Figure S2)—roughly 4 times higher than that of $^{S\text{-Me}}$γPNA1 for the same target.[34] Further increase in

temperature resulted in further increase in the rate of strand invasion, but the rate increase was less pronounced from 37 to 50 °C than from 22 to 37 °C. This result indicates that at physiological temperatures, DNA double helix is sufficiently dynamic to permit strand invasion provided that the required binding free energy can be met. Similarly, we found that the rate of strand invasion was strongly dependent on ionic strength (Figure 3B). No binding was observed at 100% physiologically simulated ionic strength (2 mM $MgCl_2$, 150 mM KCl, 10 mM sodium phosphate; pH 7.4)[40] after a 16 hr-incubation at 37 °C. The lack of productive binding, in this case, is not due to the lack accessibility—the inability of $^{R\text{-MP}}\gamma$PNA1 to gain access to the nucleobase targets, but rather due to the lack of binding free energy—as demonstrated in a recent study.[41] Once bound, the complex dissociated rather slowly, with a half-life of ~ 2 hrs (Figure 3B, *Inset*). This result indicates that it should be possible to perform *in vitro* experiments under physiologically simulated conditions with the pre-bound complex, as long as the experiments can be carried out within this time frame.

To assess the generality of $^{R\text{-MP}}\gamma$PNA binding, we determined the effects of oligo length and sequence composition on the invasion efficiency. Our results showed that although a 10mer, $^{R\text{-MP}}\gamma$PNA2a, was unable to invade dsDNA, addition of just three nucleotides ($^{R\text{-MP}}\gamma$PNA2b) restored the binding (Figure 4A, lane 3), and the efficiency gradually increased with increasing oligo length (lanes 4 to 6). However, we do not expect this trend to continue indefinitely, because, at some point, intermolecular $^{R\text{-MP}}\gamma$PNA interactions would predominate, resulting in a gradual decrease in the invasion efficiency with further increase in oligo length. We do not consider the failure of short $^{\text{MP}}\gamma$PNAs to invade dsDNA to be a detriment to most biological applications, because targeting a unique site within a mammalian genome would statistically require a recognition site of ~ 17 nucleotides in length,[42] which is within the recognition repertoire of $^{R\text{-MP}}\gamma$PNAs, and, if necessary, these oligos could be further chemically modified to provide the necessary binding free energy.[37, 41, 43]

Interestingly, we observed that unlike PNAs, even with the latest design which still requires a minimum of 40% A/T content,[33] $^{R\text{-MP}}\gamma$PNAs can invade any sequence of B-DNA, ranging from 0 to 100% G/C content (Figure 4B). Formation of these complexes occurred in a highly-sequence specific manner. Neither cross-incubation of these oligos with the other DNA targets (Figure S3), nor incubation of the G/C-rich oligos ($^{R\text{-MP}}\gamma$PNA8, $^{R\text{-MP}}\gamma$PNA8b or $^{R\text{-MP}}\gamma$PNA8bm) with DNA containing single-base mismatches (Figure S4), resulted in formation of the invasion complex. However, it should be emphasized as in the case with any intermolecular recognition event, precaution must be exercised in designing the nucleobase sequence to minimize self-hybridization because of the strong $^{R\text{-MP}}\gamma$PNA-$^{R\text{-MP}}\gamma$PNA interaction, in order to achieve optimum invasion.

In summary, we have shown that (*R*)-MiniPEG-containing γPNAs, 13 to 20 nucleotides in length, can invade any sequence of double helical B-DNA. Recognition, in this case, occurs in a highly sequence-specific manner (*vide infra*) through Watson-Crick base-pairing. The crystal structure of a $^{S\text{-Me}}\gamma$PNA-DNA duplex has been determined,[44] showing the methyl groups projecting outward toward the solvent. We do not expect the MP side-chains to behave any differently. The main advantages of MP over Me are improvements in water solubility and biocompatibility, and suppression of aggregation and nonspecific binding.

Though $^{R\text{-MP}}\gamma$PNA binding is presently limited to relatively low ionic strengths, like all other strand invading PNAs with the exception of triplex binding, this is not because they are unable to gain access to the nucleobase targets at physiological conditions (ionic strengths and temperatures), but rather because they are unable to compete with the native complementary DNA strand. This is predominantly a thermodynamic, rather than a kinetic, issue that could be resolved through molecular design. We have already demonstrated certain aspects of this design through covalent attachment of DNA intercalating agents[37] and replacement of natural nucleobases with synthetic analogues.[41] $^{R\text{-MP}}\gamma$PNAs are attractive, as compared to the other classes of antigene reagents that have been developed to date, because they are relatively easy to synthesize and they hybridize to their targets in a highly sequence specific manner in accordance with the simple rules of Watson-Crick base-pairing. Such antigene reagents could be employed in a number of biotechnology and genomic applications, including recombinant DNA, [27] genome mapping[45] and chromatin capture,[46] as well as *in vivo*, including gene regulation[28, 47] and gene correction.[48]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the National Institutes of Health (GM076251), the National Science Foundation (CHE-1012467), and the DSF Charitable Foundation.

References

[1]. Thuong NT, Helene C, Angew. Chem. Int. Ed 1993, 32, 666.
[2]. Nielsen PE, Chem. Eur. J 1997, 3, 505.
[3]. Denison C, Kodadek T, Chem. Biol 1998, 5, R129. [PubMed: 9653545]
[4]. Frank-Kamenetskii MD, Nature 1987, 328, 17. [PubMed: 3037378]
[5]. Moser HE, Dervan PB, Science 1987, 238, 645. [PubMed: 3118463]
[6]. Sun JS, Helene C, Curr. Opin. Struct. Biol 1993, 3, 345.
[7]. Giovannangeli C, Helene C, Antisense Nucleic Acid Drug Dev. 1997, 7, 413. [PubMed: 9303193]
[8]. Kielkopf CL, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC, Science 1998, 282, 111. [PubMed: 9756473]
[9]. Gottesfeld JM, Neely L, Trauger JW, Baird EE, Dervan PB, Nature 1997, 387, 202. [PubMed: 9144294]
[10]. White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Nature 1998, 391, 468. [PubMed: 9461213]
[11]. O'Hare CC, Mack D, Tandon M, Sharma SK, Lown JW, Kopka ML, Dickerson RE, Hartley JA, Proc. Nat. Acad. Sci. U.S.A 2002, 99, 72.
[12]. Greisman HA, Pabo CO, Science 1997, 275, 657. [PubMed: 9005850]
[13]. Jantz D, Amann BT, Gatto GJ, Berg JM, Chem. Rev 2004, 104, 789. [PubMed: 14871141]
[14]. Beerli RR, Barbas CFI, Nat. Biotech 2002, 20, 135.
[15]. Moore M, Klug A, Choo Y, Proc. Natl. Acad. Sci. USA 2001, 98, 1437. [PubMed: 11171969]
[16]. Garvie CW, Wolberger C, Molecular Cell 2001, 8, 937. [PubMed: 11741530]
[17]. Gowers DM, Fox KR, Nucleic Acids Res. 1999, 27, 1569. [PubMed: 10075986]
[18]. Kelly JJ, Baird EE, Dervan PB, Proc. Natl. Acad. Sci. U.S.A 1996, 93, 6981. [PubMed: 8692930]
[19]. Kielkopf CL, Baird EE, Dervan PB, Rees DC, Nat. Struct. Biol 1998, 5, 104. [PubMed: 9461074]

[20]. Nielsen PE, Egholm M, Berg RH, Buchardt O, Science 1991, 254, 1497. [PubMed: 1962210]
[21]. Nielsen PE, Acc. Chem. Res 1999, 32, 624.
[22]. Egholm M, Nielsen PE, Buchardt O, Berg RH, J. Am. Chem. Soc 1992, 114, 9677.
[23]. Nielsen PE, Christensen L, J. Am. Chem. Soc 1996, 118, 2287.
[24]. Bentin T, Nielsen PE, Biochemistry 1996, 35, 8863. [PubMed: 8688422]
[25]. Smulevitch SV, Simmons CG, Norton JC, Wise TW, Corey DR, Nat. Biotechnol 1996, 14, 1700. [PubMed: 9634855]
[26]. Zhang X, Ishihara T, Corey DR, Nucleic Acids Res. 2000, 28, 3332. [PubMed: 10954602]
[27]. Komiyama M, Aiba Y, Yamamoto Y, Sumaoka J, Nat. Protoc 2008, 3, 655. [PubMed: 18388948]
[28]. Janowski BA, Kaihatsu K, Huffman KE, Schwartz JC, Ram R, Hardy D, Mendelson CR, Corey DR, Nat. Chem. Biol 2005, 1, 210. [PubMed: 16408037]
[29]. Smolina IV, Demidov VV, Soldatenkov VA, Chasovskikh SG, Frank-Kameneteskii MD, Nucleic Acid Res. 2005, 33, e146. [PubMed: 16204449]
[30]. Bentin T, Larsen HJ, Nielsen PE, Biochemistry 2003, 42, 13987. [PubMed: 14636067]
[31]. Kaihatsu K, Shah RH, Zhao X, Corey DR, Biochemistry 2003, 42, 13996. [PubMed: 14636068]
[32]. Lohse J, Dahl O, Nielsen PE, Proc. Natl. Acad. Sci. U.S.A 1999, 96, 11804. [PubMed: 10518531]
[33]. Demidov VV, Protozanova E, Izvolsky KI, Price C, Nielsen PE, Frank-Kamenetskii MD, Proc. Natl. Acad. Sci. U.S.A 2002, 99, 5953. [PubMed: 11972051]
[34]. He G, Rapireddy S, Bahal R, Sahu B, Ly DH, J. Am. Chem. Soc 2009, 131, 12088. [PubMed: 19663424]
[35]. Harris JM, Chess RB, Nat. Rev. Drug Discov 2003, 2, 214. [PubMed: 12612647]
[36]. Sahu B, Sacui I, Rapireddy S, Zanotti KJ, Bahal R, Armitage BA, Ly DH, J. Org. Chem 2011, 76, 5614. [PubMed: 21619025]
[37]. Rapireddy S, He G, Roy S, Armitage BA, Ly DH, J. Am. Chem. Soc 2007, 129, 15596. [PubMed: 18027941]
[38]. Scholten PM, Nordheim A, Nucleic Acids Res. 1986, 14, 3981. [PubMed: 3714469]
[39]. Furlong JC, Lilley DMJ, Nucleic Acids Res. 1986, 14, 3995. [PubMed: 3012460]
[40]. Santoro SW, Joyce GF, Proc. Nat. Acad. Sci. U.S.A 1997, 94, 4262.
[41]. Rapireddy S, Bahal R, Ly DH, Biochemistry 2011, 50, 3913. [PubMed: 21476606]
[42]. Liu Z, Venkatesh SS, Maley CC, BMC Genomics 2008, 9, 509. [PubMed: 18973670]
[43]. Chenna V, Rapireddy S, Sahu B, Ausin C, Pedroso E, Ly DH, ChemBioChem 2008, 9, 2388. [PubMed: 18816545]
[44]. Yeh JI, Shivachev B, Rapireddy S, Crawford MJ, Gil RR, Du S, Madrid M, Ly DH, J. Am. Chem. Soc 2010, 132, 10717. [PubMed: 20681704]
[45]. Lin J, Lai Z, Aston C, Jing J, Anantharaman TS, Mishra B, White O, Daly MJ, Minton KW, Venter JC, Schwartz DC, Science 1999, 285, 1558. [PubMed: 10477518]
[46]. Dekker J, Rippe K, Dekker M, Kleckner N, Science 2002, 295, 1306. [PubMed: 11847345]
[47]. Mollegaard NE, Buchardt O, Egholm M, Nielsen PE, Proc. Natl. Acad. Sci. USA 1994, 91, 3892. [PubMed: 8171008]
[48]. Vasquez KM, Narayanan L, Glazer PM, Science 2000, 290, 530. [PubMed: 11039937]

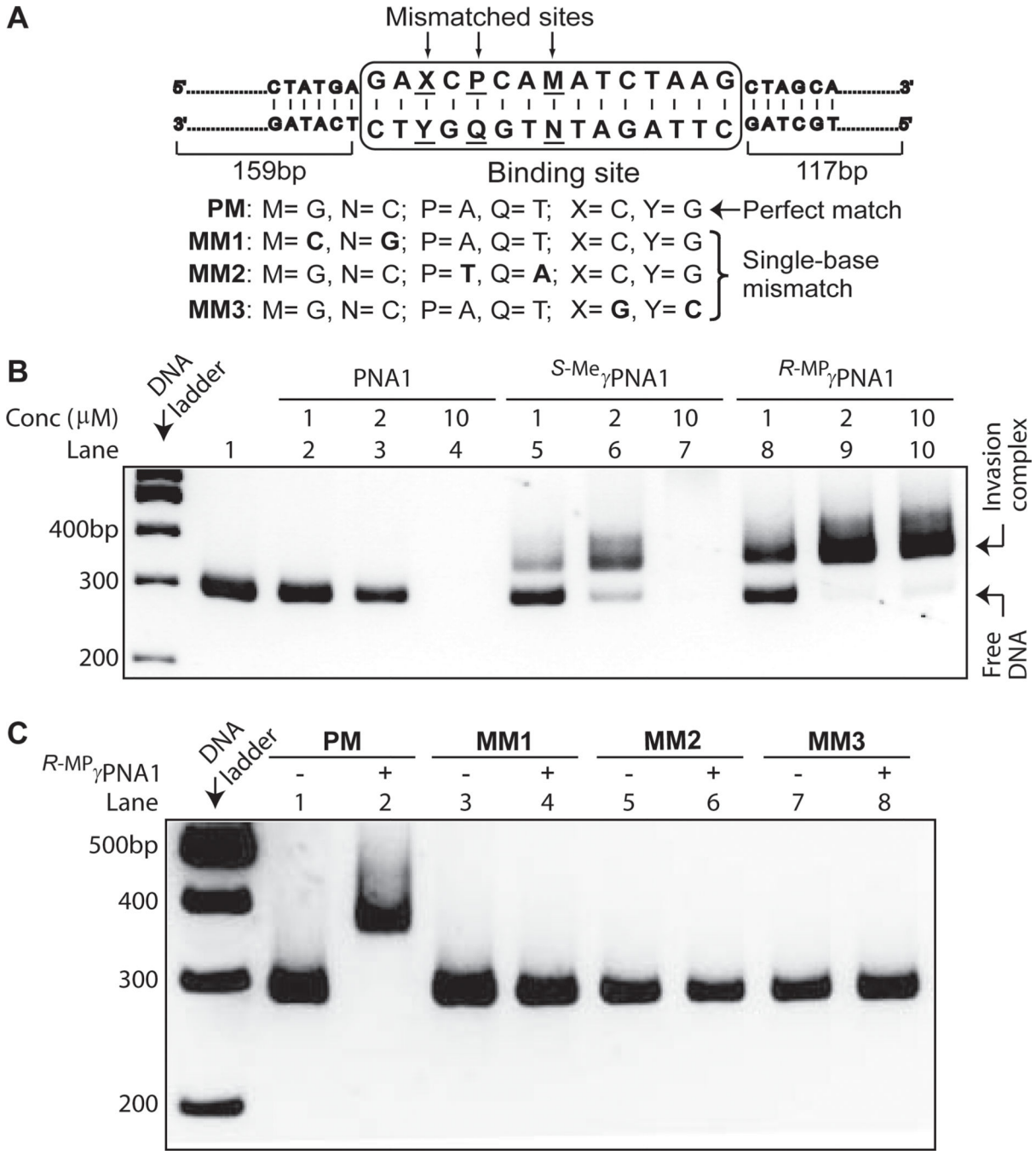


Figure 1.

Effects of backbone modification and target sequence on binding efficiency. (**A**) Schematic of the DNA targets. Results of the gel-shift assays following incubation of 0.1 μM DNA containing (**B**) **PM** binding site with different concentrations of PNA1, $^{S\text{-Me}}\gamma$PNA1, and $^{R\text{-MP}}\gamma$PNA1, and (**C**) **PM**, **MM1**, **MM2**, and **MM3** binding sites with 2 μM $^{R\text{-MP}}\gamma$PNA1 in 10 mM NaPi buffer at 37 °C for 16 hrs. The samples were separated on 10%-nondenaturing PAGE gel and stained with SYBR-Gold.

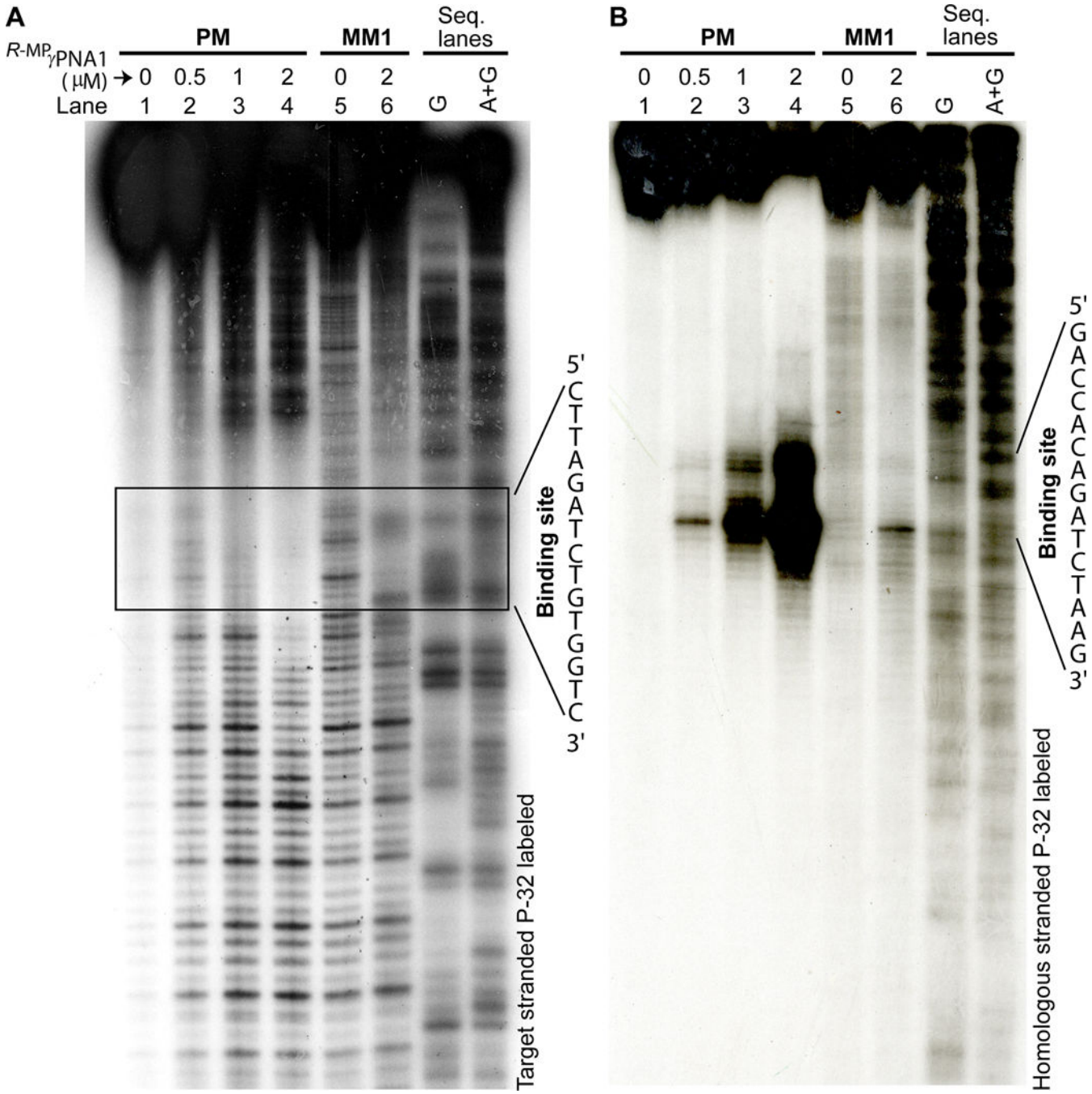


Figure 2.

Results of (**A**) DNase-I footprinting and (**B**) DEPC-chemical probing assays following incubation of 171-bp DNA fragments containing **PM** and **MM1** binding sites with different concentrations of $^{R\text{-MP}}\gamma$PNA1 in 10 mM NaPi buffer at 37 °C for 16 hrs. In (**A**) the target strand was labeled with P-32, whereas in (**B**) it was the homologous strand—both at the 3'-end.

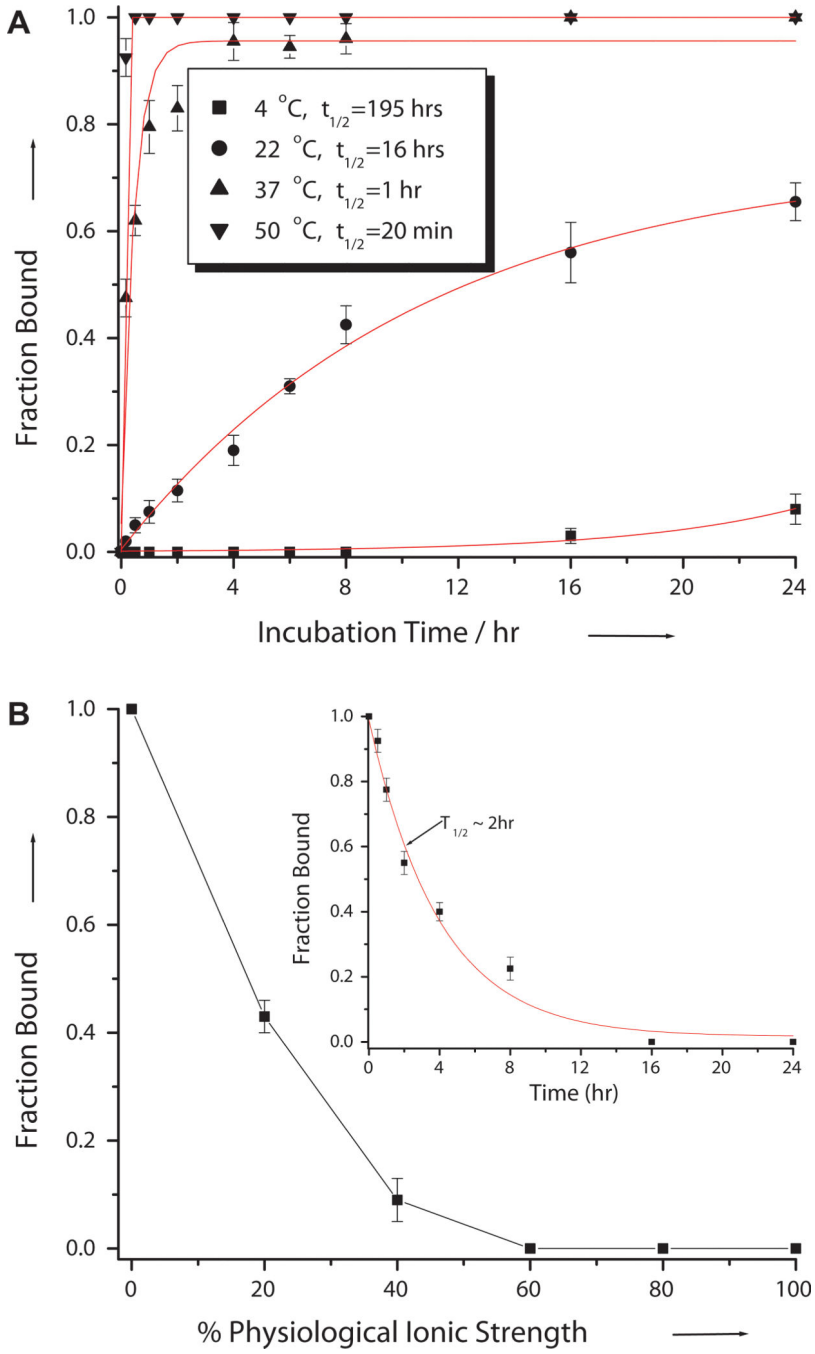


Figure 3.

Effects of (**A**) temperature and (**B**) ionic strength on the invasion efficiency. The amounts of fraction bound was determined by gel-shift assays following incubation of 0.1 μM DNA containing **PM** binding site with 2 μM of $^{R\text{-MP}}\gamma$PNA1 in (**A**) 10 mM NaPi buffer for 16 hrs at the indicated temperatures, and (**B**) different percentages of physiological ionic strength (2 mM $MgCl_2$, 150 mM KCl, 10 mM NaPi) at 37 °C for 16 hrs, followed by electrophoretic separation and SYBR-Gold staining. (**B**) *Inset:* Profile of the $^{R\text{-MP}}\gamma$PNA1•DNA complex

dissociating as a function of time after reconstituting the sample with 100% physiological ionic strength. $t_{1/2}$ is defined as the time it took to reach 50% binding.

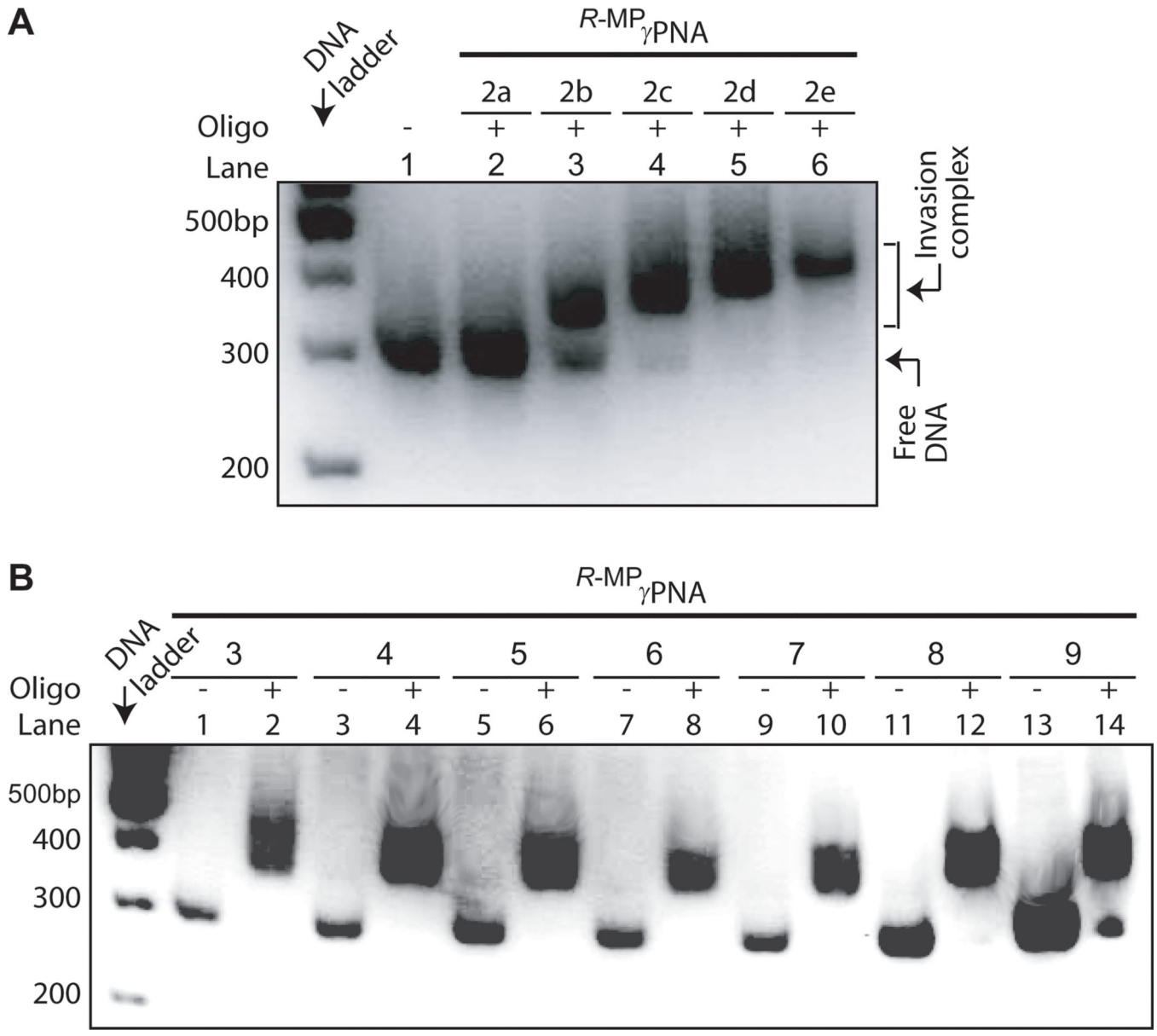
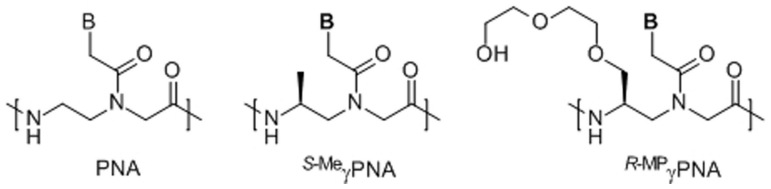


Figure 4.

Results of the gel-shift assays showing the effects of oligo (**A**) size and (**B**) sequence composition on the invasion efficiency. Gel-shift assays were performed under identical conditions as stated in the Figure 1 caption at a DNA:$^{R\text{-MP}}\gamma$PNA ratio of 20:1. The sequence of the $^{R\text{-MP}}\gamma$PNA oligos and corresponding DNA targets are shown in Table 1 and Table S1, respectively.



Scheme 1.

Chemical structures of (**A**) PNA, (**B**) *L*-alanine-derived γPNA ($^{S\text{-Me}}$γPNA), and (**C**) (*R*)-MiniPEG-containing γPNA ($^{R\text{-MP}}$γPNA). See ref. [36] for the synthesis of $^{R\text{-MP}}$γPNA monomers; the methyl ether protecting group of miniPEG side-chain is removed in the final cleavage/deprotection step of oligomer synthesis.

Table 1.

Oligonucleotides employed in this study.

Oligonucleotide	Sequence	Length	G/C [%]
PNA1	H-K-GACCACAGATCTAAG-K-NH_2	15	47
(S)-Me-γPNA1	H-K-GACCACAGATCTAAG-K-NH_2	15	47
(R)-MP-γPNA1	H-K-GACCACAGATCTAAG-K-NH_2	15	47
(R)-MP-γPNA2a	H-K-CAGATCTAAG-K-NH_2	10	40
(R)-MP-γPNA2b	H-K-CCACAGATCTAAG-K-NH_2	13	46
(R)-MP-γPNA2c[a]	H-K-GACCACAGATCTAAG-K-NH_2	15	47
(R)-MP-γPNA2d	H-K-GAGACCACAGATCTAAG-K-NH_2	17	47
(R)-MP-γPNA2e	H-K-TATGAGACCACAGATCTAAG'K-NH_2	20	40
(R)-MP-γPNA3	H-K-ATTTAATAATATAAT-K-NH_2	15	0
(R)-MP-γPNA4	H-K-CTAAACTAATGTAAT-K-NH_2	15	20
(R)-MP-γPNA5	H-K-GATTACATAGATGTC-K-NH_2	15	33
(R)-MP-γPNA6	H-K-TGCGTGAGCATCAGTK-NH_2	15	53
(R)-MP-γPNA7	H-K-CAGTGTCACGCACGG-K-NH_2	15	67
(R)-MP-γPNA8	H-K-CGGACGCAG GCTCGC K-NH_2	15	80
(R)-MP-γPNA9	H-K-CGCCCGCCGCCCGCC-K-NH_2	15	100
(R)-MP-γPNA8b	H-K-GCGAGCCTGCGTCCG-K-NH_2	15	80
(R)-MP-γPNA8bm	H-K-CGGACGCACGCTCGC-K-NH_2	15	80

[a] (R)-MP-γPNA2c and (R)-MP-γPNA1 are the same oligonucleotides, and (R)-MP-γPNAB and l-lysine. (R)-MP-γPNABb are complementary to one another; K: L-lysine.