

Segmented hybridization probes: modulating target affinity and base pairing selectivity

SIMONE EGETENMEYER, ELENA GEIGER, CLEMENS RICHERT*

Institute for Organic Chemistry, University of Stuttgart, Stuttgart, Germany

* Corresponding author: lehrstuhl-2@oc.uni-stuttgart.de

Abstract

Short hybridization probes bind their targets with greater base pairing fidelity, but with lower affinity than longer probes. Furthermore, their target sequence is shorter, and thus more likely not to be unique in a given genome. Long hybridization probes provide increased affinity, and their sequences are more unique, but their duplexes tolerate mismatches more readily, without a significant depression in melting point. It was reasoned that segmenting longer hybridization probes by introducing flexible, abasic linkers might lead to oligonucleotides that retain some of the sequence selectivity of short probes without losing too much of the target affinity of their unsegmented counterparts. A model study led to 1,3-propanediol-phosphates as linker residues. These spacer residues were introduced at different positions of hybridization probes 8-20 residues in length and their hybridization properties were studied in UV-melting curves with RNA or DNA target strands. Increases in base pairing selectivity ($\Delta\Delta T_m$ of up to -7.4°C for a single mismatch) and decreased target affinities (ΔT_m between -15 and -25°C) were found for the segmented probes when compared to their unsegmented counterparts, and so was a decreased selectivity for insertions at the site of the linker. Also, the increases in selectivity are not uniform in their magnitude and depend on sequence context and position. A favorable case appears to be a hybridization probe that contains two spacers, with one octamer as core segment, flanked by a heptamer and a pentamer as terminal segments.

Key words: oligonucleotides, DNA, RNA, hybridization, base pairing

Introduction

Hybridization of synthetic oligonucleotide to a complementary sequence in a target strand is a fundamental process in genetics and diagnostics. Southern and Northern blots are important methods in molecular biology for detecting DNA or RNA sequences. The sequence-specific binding of primers to sections of genomic DNA is critical for amplification by the polymerase chain reaction. The process of hybridization is governed by the base pairing rules of Watson and Crick (Watson, 1953). The advent of microarrays (Fodor, 1991) has led to a new level of "economy of scale" in hybridization, with multiple hybridization events being induced simultaneously in the same system consisting of a surface with probes and a solution with multiple targets (Lipshutz, 1999). Massively parallel hybridization increases the extent of cross-hybridization, though, as duplexes with only partial complementarity form in ever-increasing number, causing false positives and a strong background signal (Marshall, 2004; Plutowski, 2005; Siegmund, 2008).

Nature evolved DNA as storage material for genetic information, not as material for hybridization probes, explaining why sequence fidelity in hybridization is rather limited. To overcome the natural limits, modified oligonucleotide structures are being developed, and probe selection is aided by computational methods (Gambin, 2011). Better known approaches for modified oligonucleotides include locking the backbone of hybridization probes into a conformation preorganized toward binding the fully complementary target (Singh, 1998; Koshkin, 1998; Prestinari, 2011), appending ligands that help to interrogate the target (Maltseva et al., 1994; Dogan et al., 2004; Narayanan, 2004; Patra, 2009), or adjusting for differences in G/C content (Ahlborn, 2007). Structured probes and peptide nucleic acids have also been used to improve selectivity in hybridization (Bonnet, 1999; Kushon et al., 2001). In an extension of the work on structured probes, a recent publication proposes toehold exchange hybridization as a means to optimize hybridization systems (Zhang, 2012), however

discrimination factors as low as 3 are still found and it is unclear whether the approach, which requires twice as many oligonucleotides and is currently based on gels, will be useful for microarrays.

A general phenomenon in hybridization is that shorter probes give larger changes in UV-melting point for single mismatches as compared to longer probes. For very short probes, such as hexamers or octamers, a single mismatch can lead to a drop in melting point of 20 °C or more, a value that is rarely found for longer hybridization probes, even if the mismatch is at an ideal position, in the interior of the sequence. While a single mismatch is a near-catastrophic event for the short probe, affecting a large percentage of its base pairs, the gradual breakdown of cooperativity makes the increasingly smaller percentage of base pairs lost less catastrophic for long duplexes. For very long or genomic DNA, individual mismatches are inconsequential for duplex stability. There is a trade-off between choosing the longest possible sequence in order to be statistically unique in a given genome and choosing the shortest possible probe to avoid false positives. The typical compromise used for synthetic microarrays and many primers is the length between 16 nucleotides (the length required to be statistically unique in the human genome) and 25 nucleotides. This length range also encompasses the number of nucleotides found in microRNAs, whose selective detection in clinical samples is a diagnostic challenge (Wark, 2008). It has been previously reported that full length probes perform poorly in detecting microRNAs differing by just one nucleotide (Egetenmeyer, 2011), prompting the question whether segmenting longer hybridization probes into a series of shorter hybridization probes would lead to a better compromise between affinity and selectivity. Here we report the results from an experimental study into this issue.

Materials and methods

Sequences are given in 5' to 3' direction. Unmodified RNA was obtained from Biomers Inc. (Ulm, Germany) in cartridge-purified quality. The segmented DNA probe strands were synthesized via automated synthesis with the use of "spacer phosphoramidite 18" (**1**), "spacer phosphoramidite 9" (**2**), both received from Glen Research (Sterling, VA, USA), or 1,3-propanediol spacer phosphoramidite (**3**) synthesized following a literature-

known route (Seela, 1987). The DNA syntheses were carried out on a Perseptive Biosystems 8909 Expedite DNA synthesizer, following the protocol provided by the manufacturer for 1 μ M scale. The DNA synthesis reagents and controlled pore glass (CPG) were from Prologo (Hamburg, Germany). For cleavage and deprotection of the DNA, the CPG was treated with aqueous ammonia (27%, 1 ml) and heated to 60 °C for 2 h. The oligonucleotides were purified using C18 cartridges (Sep-Pak Vac 3cc) from Waters (Milford, MA, USA) with triethylammonium acetate buffer (0.1 M, pH 7) and a step gradient of CH₃CN. Molecular caps U_{AQ} and ogOA, shown in Figures 4 and 5, were introduced following the methods described earlier (Narayanan, 2004; Egetenmeyer, 2011). The analytical data and MALDI-TOF mass spectra of modified oligonucleotides are provided in the supporting information.

UV-melting experiments

UV-melting curves were measured at 260 nm with a Lambda 25 spectrophotometer equipped with a PTP 6-controlled Peltier-heated six-cell holder (Perkin Elmer, Waltham, MA, USA). DNA and RNA strands (1 μ M) were dissolved in sodium phosphate buffer (10 mM, 0.6 ml) containing NaCl (1 M). For each sample, four curves were detected in the temperature range of 5 °C to 80 °C with a heating or cooling rate of 1 °C/min. Melting points were determined using the Lambda 2.0 software of the spectrophotometer, and the values are given as average of the four curves. Representative melting curves are shown in the supporting information.

Results and discussion

Our experimental approach had three stages: synthesis of a linker building block, automated DNA synthesis, and UV-melting analysis of duplexes with target strands. The criteria for evaluating the effect of segmentation on target binding were (i) the decrease in the UV-melting point for the fully complementary sequence when one or two linkers were being introduced (affinity), and (ii) the change in melting point depression for a single mismatch at a predefined position (selectivity). We assumed that by properly choosing the length and sequence of the segments and the chemical structure of the linker(s), we would be able to avoid biphasic or triphasic transition caused by partial duplex intermediates (Fig. 1). We hoped to retain the cooperative

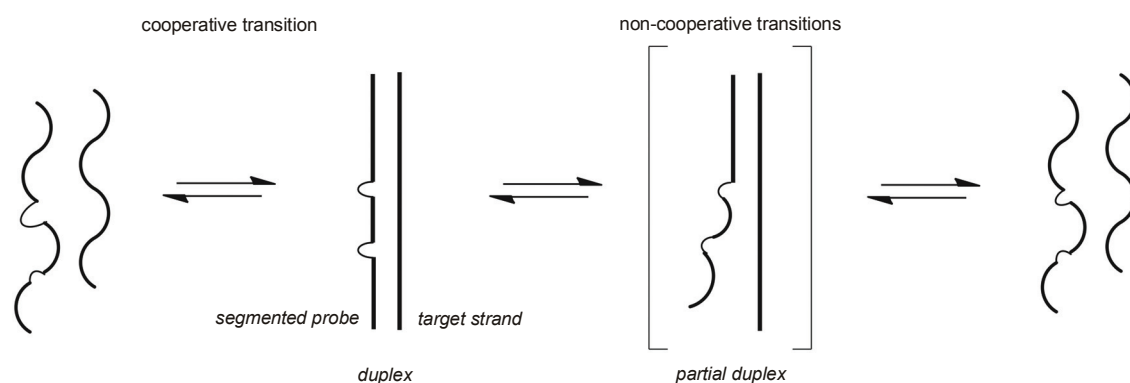


Fig. 1. Melting transitions for a duplex between a segmented probe and an RNA target strand containing a single mismatched base. A cooperative transition, avoiding significant population of partial duplex states (left-hand equilibrium) is desirable for hybridization probes. A step-wise, non-cooperative transition (right-hand equilibrium) would make segments behave like individual probes, resulting in broad transitions that complicate the highly selective detection of target sequences, particularly when performed in massively parallel fashion

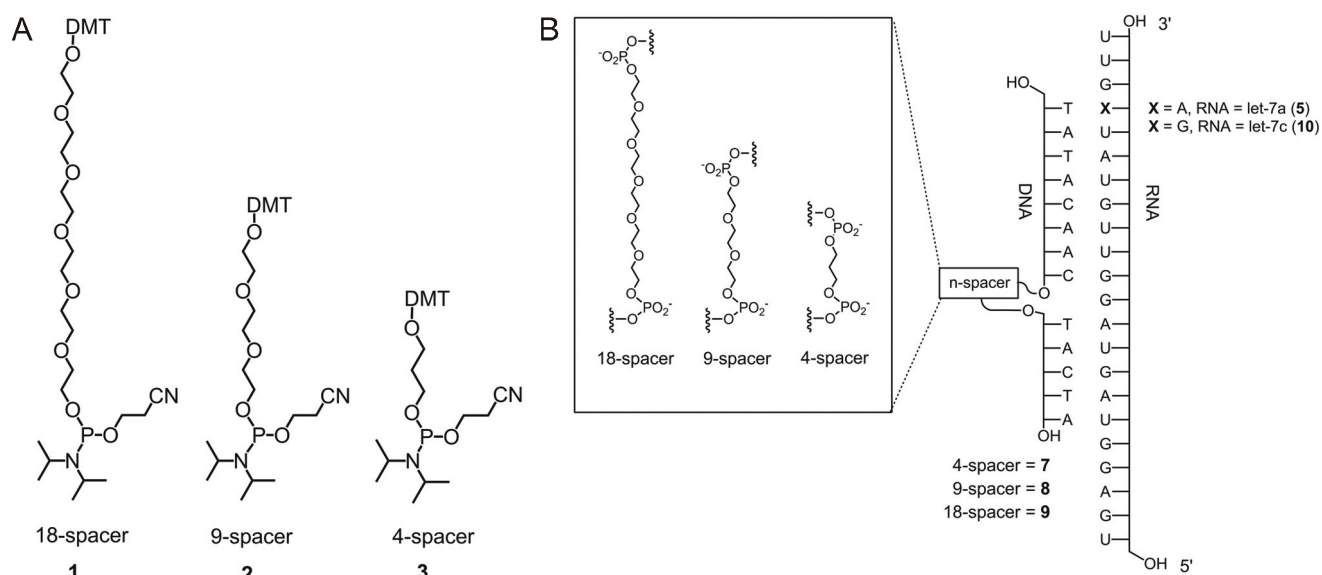


Fig. 2. A) Spacer phosphoramidites employed; B) strands of the sequence system that was studied by UV-melting to determine an appropriate spacer length

transition between just two states: full duplex and single strands.

The sequences shown in Figure 2 were used to find an appropriate linker (Fig. 2). Linkers consisting of 4-, 9-, or 18-membered chains were chosen as spacers between segments of the probes. Phosphoramidites **1** and **2** are commercially available, whereas phosphoramidite **3** was synthesized following a previously published protocol (Seela, 1987).

Members of the let-7 family of microRNAs (Abbott et al., 2005) were selected as target strands. The results of melting point experiments are shown in Table 1.

The spacer structure was found to significantly affect the stability of the duplexes, as well as the selectivity of binding a complementary strand with or without a mismatch. Compared to the duplex of unsegmented probe **4**, the duplex with the octamer probe **6** with target **5**, showed a drop in T_m by almost 50°C , but an increase in the melting point difference between perfect match (PM) and mismatch-containing sequence (MM). All segmented probes **7-9** consisting of the octamer plus a pentamer, linked by the spacer, gave higher melting points than octamer **6**. At the same time, all depicted a larger ΔT_m value for the mismatch-containing duplex versus the per-

Table 1. UV-melting points of duplexes of segmented or unsegmented DNA probe strands and miRNA let-7a or let-7c as target strands. Conditions: probe strand (1.5 μM) and RNA target strand (1.5 μM) in sodium phosphate buffer (10 mM), pH 7, containing 1 M NaCl

DNA probe	RNA target let-7a (5) T_m [$^{\circ}\text{C}$]	RNA target let-7c (10) T_m [$^{\circ}\text{C}$]	$\Delta T_m^{[a]}$ [$^{\circ}\text{C}$]
<i>d</i> (AACTATACAACCTACTACCTCA) 4	70.2	68.4	-1.8
<i>d</i> (TATACAAC) 6	20.5	16.6	-3.9
<i>d</i> (TATACAAC 4-spacer TACTA) 7	34.6	30.2	-4.4
<i>d</i> (TATACAAC 9-spacer TACTA) 8	32.5	29.9	-2.6
<i>d</i> (TATACAAC 18-spacer TACTA) 9	27.7	24.7	-3.0

[a] Difference in melting point of the duplexes with let-7c and let-7a, the mismatch-containing and perfectly matched duplex, respectively

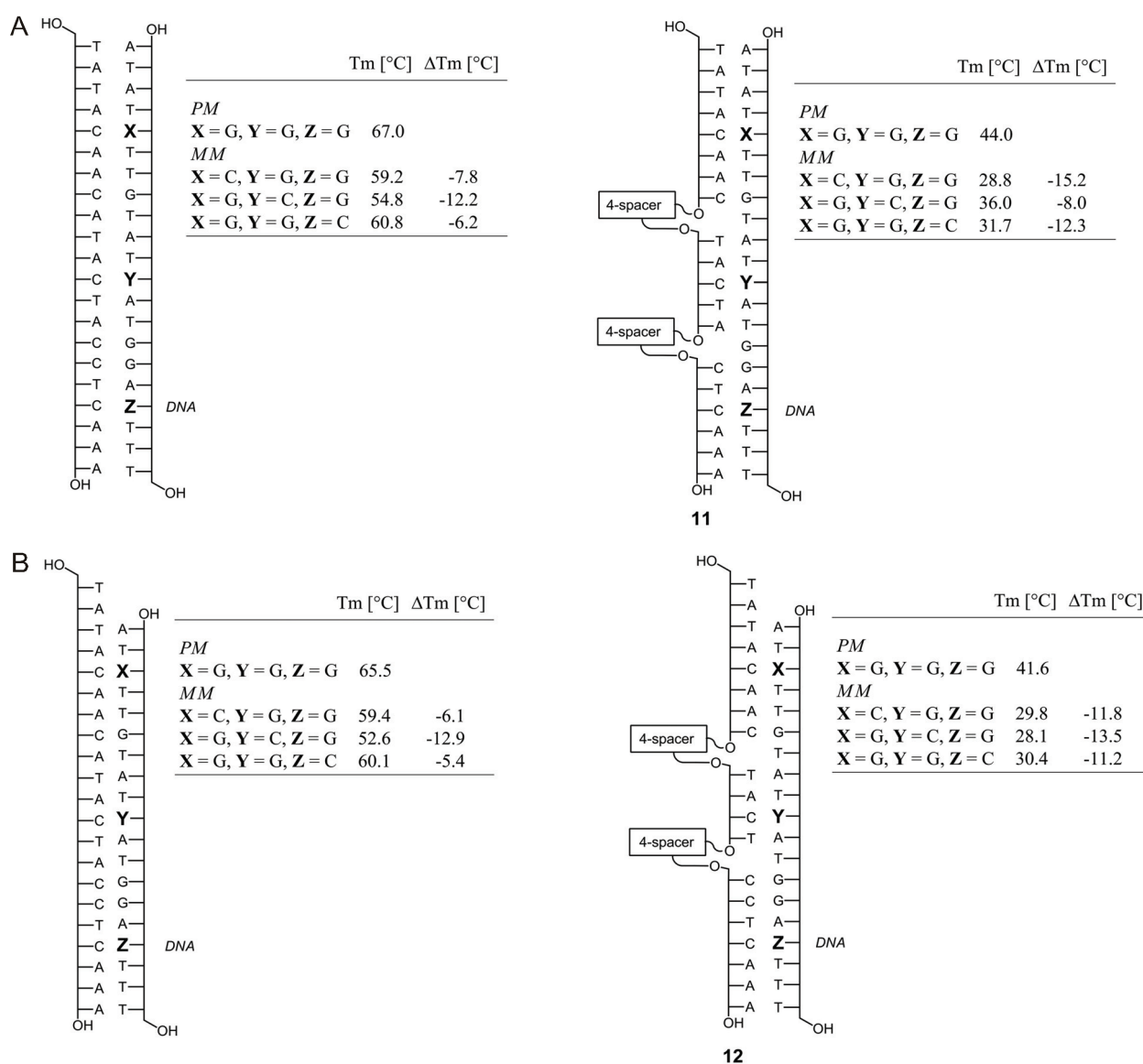


Fig. 3. Sequences of duplexes without (control) or with two spacers in the backbone of the probe strand and target strands containing none or a single mismatch and their melting points. The UV-melting points of the duplexes are shown next to each duplex for ease of reading. The abbreviation “PM” stands for perfectly matched duplexes, whereas “MM” denotes sequences that contain a mismatch. Conditions for UV-melting experiments: probe and target strands (1.5 μM) in sodium phosphate buffer (10 mM) containing 1 M NaCl

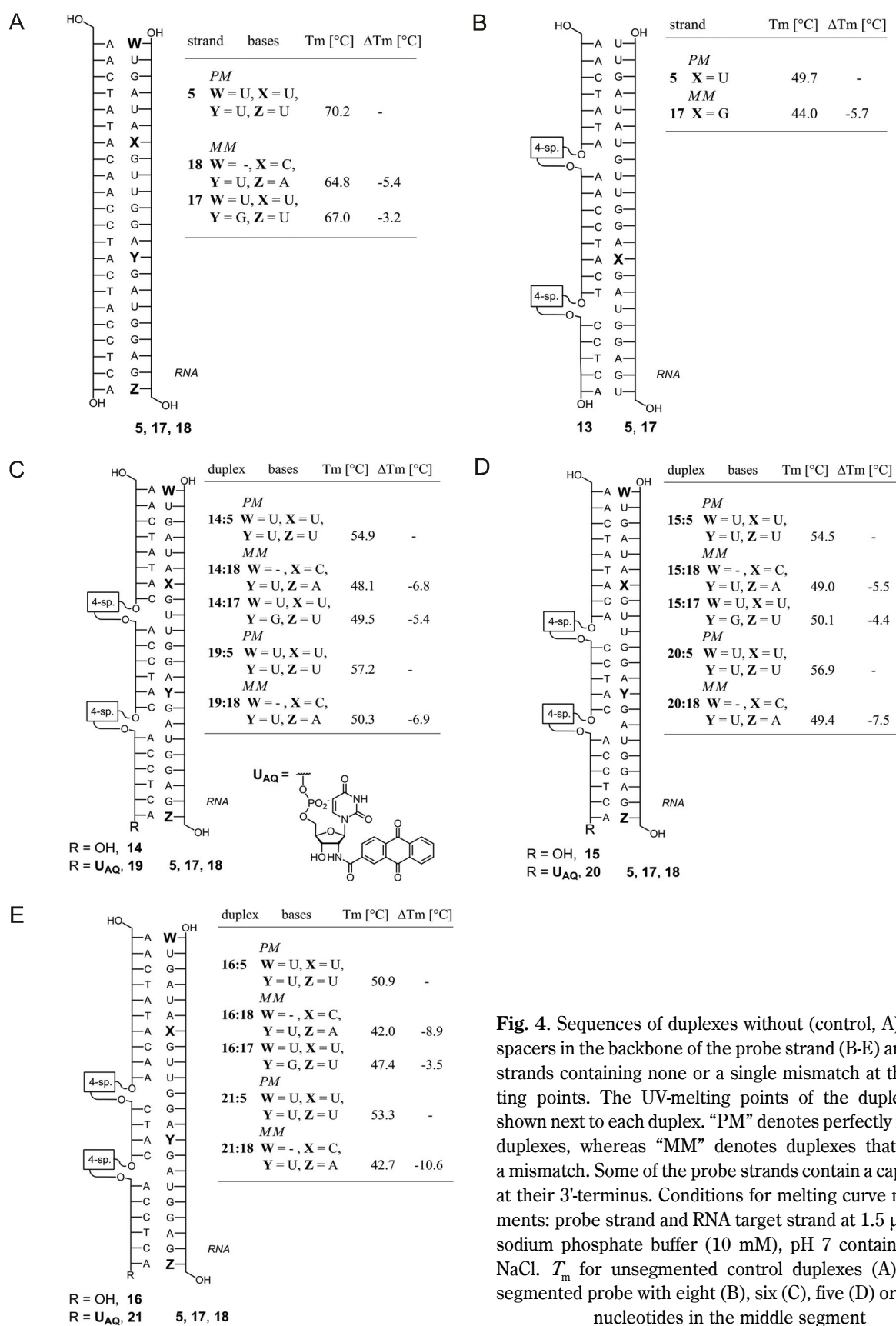


Fig. 4. Sequences of duplexes without (control, A) or with spacers in the backbone of the probe strand (B-E) and target strands containing none or a single mismatch at their melting points. The UV-melting points of the duplexes are shown next to each duplex. “PM” denotes perfectly matched duplexes, whereas “MM” denotes duplexes that contain a mismatch. Some of the probe strands contain a cap (U_{AQ}) at their 3'-terminus. Conditions for melting curve measurements: probe strand and RNA target strand at 1.5 μ M each, sodium phosphate buffer (10 mM), pH 7 containing 1 M NaCl. T_m for unsegmented control duplexes (A), T_m for segmented probe with eight (B), six (C), five (D) or four (E) nucleotides in the middle segment

fectly matched duplex than full length 22mer **4**. The 1,3-propanediol spacer led to slightly better selectivity for **5** over **10**, with a ΔT_m value of -4.4°C compared to -3.9°C for the octamer. The absolute T_m of the duplex with this spacer went up by $+14.1^\circ\text{C}$ compared to that of octamer **6** with let-7a. As a consequence, the 4-spacer with a 1,3-propanediol phosphate linker was chosen as most suitable for subsequent work with segmented probes.

The next step was to study DNA:DNA duplexes with doubly segmented probes. Figure 3 shows the sequences of the duplexes. For clarity, the melting points are shown adjacent to them. The unsegmented control probe or segmented probe **11** were hybridized to target strands that were either fully complementary or contained a single mismatch at one of three different positions (X, Y, and Z). Segmented probe **11** showed lower affinity, but gave better discrimination against single C:C mismatches at the two distal mismatch positions, when compared to the unsegmented probe (Fig. 3a). At the central position, it performed more poorly than the control probe. Seemingly minor changes in the length of the target and the segments in probe **12** resulted in a duplex with improved ΔT_m values for every of the mismatches studied (Fig. 3b).

The segments of the probe strand were then varied to improve the selectivity in binding miRNA let-7a over let-7e (**17**). Sequences and melting points of the duplexes are given in Figure 4. The duplexes of the probe strands with let-7e contain a A:G mismatch in the middle region. Probe strands with a longer segment in the center region that faces the mismatch site showed improved selectivity compared to the unsegmented probe. The latter gave a ΔT_m of -3.2°C for the center (Fig. 4a), compared to -5.7°C for the segmented probe strand with an octamer center (Fig. 4b). Among the segmented probes, the one with a hexamer core gave the highest T_m for the perfectly matched duplex, and this T_m could be further increased by appending a molecular cap at the 3'-terminus (probe **19**, Fig. 4c). The U_{AQ} cap was introduced as reported earlier (Patra, 2009). It helped to improve the mismatch discrimination at the 3'-terminus. A CPG for probe synthesis with the U_{AQ} cap is commercially available (Richert, 2011).

The discrimination of miRNA let-7a over let-7d (**18**) is challenging because the duplex of the probe strand employed and let-7d contained only two mismatches.

One A:A mismatch is at the terminus, and the other, internal A:C mismatch, is close to the opposite terminus. Since absolute T_m values of the duplexes decrease as the length of the terminal segments increases, one has to strike a balance to avoid unacceptably low selectivity in the center region or unfavourably low affinity. The position of the spacers in the probe was varied so that the terminal segment that contains the A:C mismatch was varied in length. Again, the trend was discernible that longer segments lead to better mismatch discrimination in the region they bind (Fig. 4d and 4e).

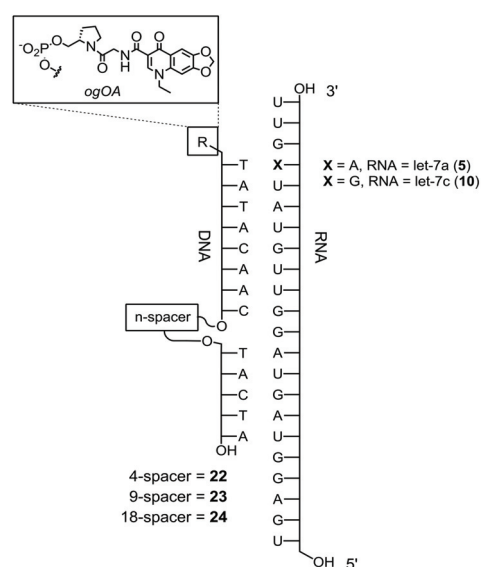


Fig. 5. Sequences of duplexes with a shortened probe featuring the *ogOA* cap with spacers of different lengths in its backbone and target strands containing none or a single mismatch. See Table 2 for the melting points

Because high fidelity recognition of miRNA let-7a over let-7c appeared rather difficult due to the single mismatch in the terminal region, a shorter probe with just one spacer and a molecular cap optimized for high fidelity recognition at the 5'-terminus was tested next. This was the *ogOA* cap that was reported recently (Egetenmeyer, 2011). It consists of an oxolinic acid residue as terminal stacking moiety, linked to the DNA chain *via* a glycine residue and a prolinol residue forming the phosphodiester linker to the terminus of the probe (Fig. 5). With the 5'-cap, the absolute T_m values for all duplexes of segmented probes were increased and the selectivity for let-7a was improved (Table 2). In the melting curve experiments, ΔT_m values in the range

Table 2. UV-melting points of duplexes of shortened segmented DNA probes with a molecular cap at their 5'-terminus, and miRNAs let-7a or let-7c as targets. Conditions are same as mentioned in Table 1

DNA probe	let-7a (5) T_m [°C]	let-7c (10) T_m [°C]	ΔT_m [°C] ^[a]
<i>ogOA-d</i> (TATACAAC 4-spacer TACTA) 22	41.7	35.8	-5.9
<i>ogOA-d</i> (TATACAAC 9-spacer TACTA) 23	39.6	33.3	-6.3
<i>ogOA-d</i> (TATACAAC 18-spacer TACTA) 24	38.4	32.0	-6.4

[a] Difference in melting point (T_m) of the duplex with let-7c, compared to the corresponding perfectly matched duplex with target strand let-7a

Table 3. UV-melting points of DNA duplexes with segmented probe strands and control duplexes, respectively. DNA target strands with or without one insertion or one deletion facing the spacer were tested. Conditions: same as those given in Table 1

DNA probe	DNA target	T_m [°C]	ΔT_m [°C] ^[a]
11	<i>d</i> (TTTGAGGTAGTATGTTGTA)	41.6	-
11	<i>d</i> (TTTGAGGTAGTATTGTTGTA)	39.2	-2.4
11	<i>d</i> (TTTGAGGTAGTA GTTGTA)	45.8	+4.2
<i>d</i> (TATACAACATACTACCTCAAA)	<i>d</i> (TTTGAGGTAGTATGTTGTA)	65.5	-
<i>d</i> (TATACAACATACTACCTCAAA)	<i>d</i> (TTTGAGGTAGTATTGTTGTA)	60.9	-4.6
<i>d</i> (TATACAACATACTACCTCAAA)	<i>d</i> (TTTGAGGTAGTA GTTGTA)	57.4	-8.1

[a] Difference in melting point compared to the corresponding perfectly matched duplex

of -5.9 to -6.4°C for the terminal mismatch were observed with different spacer lengths tested.

Finally, UV-melting curve experiments were carried out for duplexes with segmented probe strands and target strands with an additional thymidine residue (insertion) or a missing nucleotide (deletion) in the region facing the spacer (Table 3, Fig. 6). To minimize the cost, these experiments were performed with DNA target strands. The target strand with the insertion was recognized by the segmented probe with a small decrease in T_m (-2.4°C), but the continuous probe gave a somewhat larger decrease in T_m (-4.6°C). The deletion in the target strand led to an *increase* in T_m for the duplex of the segmented probe and the expected *decrease* for the unsegmented probe, with a T_m of -8.1°C. The segmented probe, therefore, does not appear to be suitable for discrimination against target strands containing deletions at the sequence region where the spacer is located, at least when the spacer is a flexible chain and the target is DNA.

Conclusions

Hybridization of oligonucleotide probes and the dissociation of duplexes into single strands upon heating are

highly cooperative processes (Saenger, 1984). Sigmoidal transitions are found and the data are usually interpreted on the basis of a two-state model (Breslauer, 1995). When the nucleation region forms a duplex, the remainder of the sequence quickly zippers up to form an entire helix. Current study explores the limits of this concept, using segments of a given oligonucleotide rather than just individual base pairs as subunits that can either act cooperatively in binding a target or not. Appearance of the melting transitions of duplexes with a doubly segmented probe is surprisingly cooperative in appearance (Figures S1, S2, Supporting Information). Furthermore, the increases in melting point over that observed for the longest segment alone are substantial, again attesting to the remaining cooperativity despite interrupting spacers. Finally, the increase in base pairing selectivity, as measured in $\Delta\Delta T_m$ values for the perfectly matched versus mismatch-containing duplexes, can be significant for segmented probes. Short segments in the center are unfavorable, though, and increases in selectivity with optimized segments come at a price. Insertions and deletions in the target sequence may not be well discriminated against the loci of the spacers. Therefore, the practical usefulness of the segmentation concept

- Richert C. (2011) *A 3'-Cap for improved target affinity and specificity*. Glen Research Report 21: 4-5.
- Saenger W. (1984) *Principles of Nucleic Acid Structure*, Springer: New York.
- Seela F., Kaiser K. (1987) *Oligodeoxyribonucleotides containing 1,3-propanediol as nucleoside substitute*. Nucl. Acids Res. 15: 3113-3129.
- Siegmund K., Ahlborn C., Richert C. (2008) *ChipCheckII – Predicting binding curves for multiple analyte strands on small DNA microarrays*. Nucleos. Nucleot. Nucl. Acids 27: 376-388.
- Singh S.K., Nielsen P., Koshkin A.A., Wengel J., (1998) *LNA (locked nucleic acids) synthesis and high affinity nucleic acid recognition*. Chem. Comm. 4: 455-456.
- Wark A.W., Lee H.J., Corn R.M. (2008) *Multiplexed detection methods for profiling microRNA expression in biological samples*. Angew. Chem. Int. Ed. Engl. 47: 644-652.
- Watson J.D., Crick F.H. (1953a) *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*. Nature 171: 737-738.
- Watson J.D., Crick F.H. (1953b) *Genetical implications of the structure of deoxyribonucleic acid*. Nature 171: 964-967.
- Zhang D.Y., Chen S.X., Yin P. (2012) *Optimizing the specificity of nucleic acid hybridization*. Nat. Chem. 4: 208-214.