# UvA-DARE (Digital Academic Repository)

## Probing alternative foldings of the HIV-1 leader RNA by antisense oligonucleotide scanning arrays

Ooms, M.P.; Verhoef, K.; Southern, E.; Huthoff, H.; Berkhout, B.

[Link to publication]

# Probing alternative foldings of the HIV-1 leader RNA by antisense oligonucleotide scanning arrays

**Marcel Ooms, Koen Verhoef[1], Edwin Southern[1], Hendrik Huthoff and Ben Berkhout***

Department of Human Retrovirology, Academic Medical Center, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands and [1]Department of Biochemistry, University of Oxford, Oxford, UK

## ABSTRACT

**Scanning arrays of antisense DNA oligonucleotides provide a novel and systematic means to study structural features within an RNA molecule. We used this approach to probe the structure of the untranslated leader of the human immunodeficiency virus type 1 (HIV-1) RNA genome. This 335 nt RNA encodes multiple important replication signals and adopts two mutually exclusive conformations. The poly(A) and the dimer initiation signal (DIS) sequences of the leader RNA are base-paired in the long-distance interaction (LDI) conformation, but both domains form distinct hairpins in the branched multiple hairpins (BMH) conformation. An RNA switch mechanism has been proposed to regulate the activity of the DIS dimerization signal that is masked in one, yet exposed in the other conformation. The two RNA conformations demonstrate discrete differences in the array-based hybridization patterns. LDI shows increased hybridization in the poly(A) region and decreased hybridization in the DIS region when compared with BMH. These results provide additional evidence for the structure models of the two alternative leader RNA conformations. We also found a correlation between the efficiency of oligonucleotide hybridization and the accessibility of the RNA structure as determined by chemical and enzymatic probing in previous studies. The array approach therefore provides a very sensitive method to detect structural differences in related transcripts.**

## INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) encodes nine proteins from the single-stranded RNA genome (Fig. 1A). Surprisingly, the most conserved part of this viral genome is not found in one of the open reading frames (ORFs), but in the 5′-untranslated leader region of 335 nt. The HIV-1 leader RNA contains multiple sequence and structure motifs that play important roles in distinct steps of the viral replication cycle (1). RNA elements within the leader that serve to control gene expression include the TAR hairpin that mediates transcriptional activation by the viral Tat protein, the poly(A) hairpin that suppresses the polyadenylation signal (2), the major splice donor and the start codon of the Gag ORF. Packaging of a dimeric form of the RNA genome in virus particles is also controlled by sequences in the leader; the dimer initiation signal (DIS) hairpin and the Ψ signal. Reverse transcription of the viral RNA genome is strictly controlled by sequence and structure elements in the central part of the leader, i.e. the primer-binding site (PBS) and accessory sequence motifs such as the primer activation signal (PAS) (3). Some of these regulatory leader signals, such as the Ψ signal, have been difficult to map in detail by mutational analysis. One reason for this is the poor understanding of the secondary structure and particularly the tertiary structure of this RNA molecule. The description of these signals is complicated further by the existence of alternative RNA conformations and overlap of critical sequence and structure elements (1).

A static branched multiple hairpins (BMH) structure of the HIV-1 leader RNA was originally proposed, but we recently documented that this RNA will adopt a more stable, alternative conformation once sufficient downstream leader sequences are included in the transcript (4). Specifically, we demonstrated that sequences of the upstream poly(A) hairpin and the downstream DIS hairpin form a long-distance interaction (LDI; Fig. 1B). This ground state structure refolds into the BMH conformation in the presence of the viral nucleocapsid protein (4). This RNA switch provides a mechanism to regulate the process of RNA dimerization because the DIS hairpin is masked in the LDI conformation, and therefore not able to dimerize efficiently. Furthermore, structural changes in the leader RNA can also have an effect on reverse transcription (5). This suggests that HIV-1 may coordinate multiple leader RNA functions through the same RNA switch, which has led us to propose the concept of an RNA checkpoint that coordinates different replication steps (5). Inspection of the complete leader RNA revealed that the AUG start codon of the Gag ORF also differentially contributes to the LDI and BMH conformations (6). This recent finding suggests that the leader RNA switch also influences HIV-1 translation.

Considerable evidence has been presented in support of the alternative conformations and their regulatory role in RNA dimerization both in HIV-1 (4,7,8) and in HIV-2 (9,10). A detailed thermodynamic pathway for the multiple structural rearrangements from LDI to BMH to RNA dimers has recently
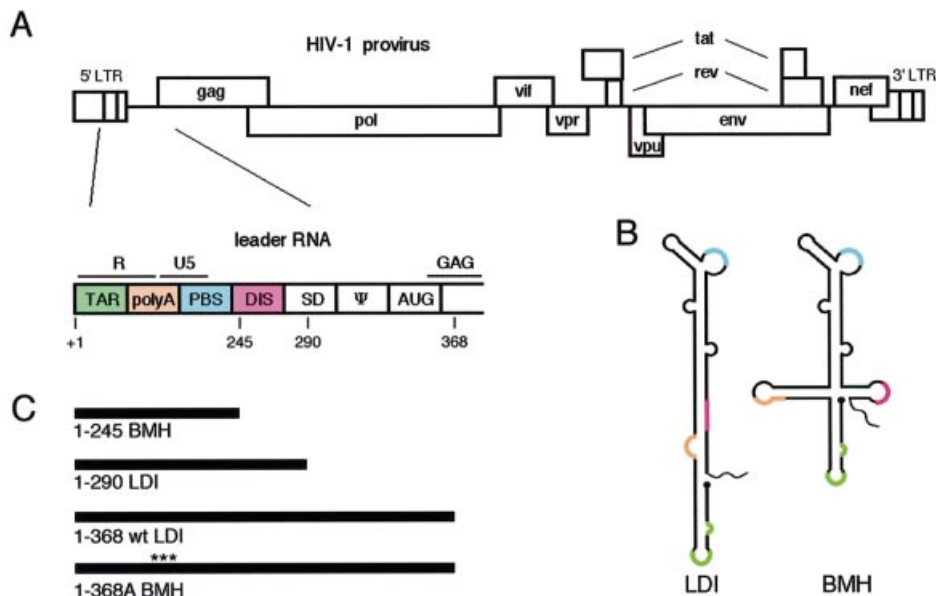
---

**Figure 1.** Alternative foldings of the HIV-1 leader RNA. (**A**) Schematic of the 9.2 kb HIV-1 genome. The 5′- and 3′-LTRs and all nine ORFs are indicated. The untranslated leader RNA consists of several regulatory domains (see text). (**B**) The HIV-1 leader RNA can adopt two conformations. The poly(A) (orange) and DIS sequences (pink) are base-paired to form a long-distance interaction in the LDI structure. The same sequences form the poly(A) and DIS hairpins in the BMH structure. More detailed base pairing schemes of several domains are shown in Figures 4 and 6. (**C**) The leader RNA transcripts used in this study start at position +1. The 1–245 transcript lacks the DIS sequence and therefore adopts the BMH structure. The DIS sequence is included in the 1–290 LDI transcript, resulting in the formation of the LDI. The 1–368 wild-type (wt) transcript adopts the LDI structure. Three stabilizing mutations were introduced in the poly(A) hairpin of the A mutant (indicated by asterisks, see Fig. 4 for details) such that the BMH structure is favored over the LDI structure.

been proposed (11). In this report, we set out to probe the alternative HIV-1 RNA conformations with antisense oligo-nucleotides. Specifically, we used scanning arrays with tethered antisense DNA oligonucleotides, ranging from 5 to 25mers, that target all possible sequences within the HIV-1 leader RNA. The use of such arrays provides an empirical approach for selecting the optimal antisense oligonucleotide for inhibition of gene expression (12–14). This novel and powerful technology is also able to detect structural features within an RNA molecule (5,15,16). The beauty of the differential hybridization experiments performed in this study is that they eliminate the influence of other parameters of heteroduplex formation such as heteroduplex stability and oligonucleotide intra- and inter-molecular base pairing, as these parameters are exactly the same for both targets in the hybridization reaction. We demonstrate that this technique is very sensitive to structural changes and thus particularly suited to detect alternative conformations of the HIV-1 leader RNA.

## MATERIALS AND METHODS

### HIV-1 templates for PCR and T7 transcription

Wild-type and mutant plasmids were used for PCR amplification and subsequent *in vitro* transcription. Plasmid Blue-5′-LTR (17) contains an XbaI–ClaI fragment of the HIV-1 subtype B molecular clone LAI (18). This HIV-1 fragment contains the complete 5′-long terminal repeat (LTR), PBS, leader and the 5′ end of the *gag* gene (positions –454 to +376 relative to the transcriptional start site +1) and was cloned into

pBluescript KS+ (Stratagene). The construction of the A mutation in the poly(A) hairpin has been described previously (17,19,20). The wild-type and mutant plasmids were used as templates for PCR amplification. The T7 promoter sequence was introduced into the PCR product with sense primers containing the T7 promoter sequence. The T7 promoter sequence was inserted immediately upstream of position +1 with the sense primer T7-2 (5′-CTA ATA CGA CTC ACT ATA GGG TCT CTC TGG TTA GAC CAG-3′). The oligonucleotides 245/226 (5′-GAG TCC TGC GTC GAG AGA GC-3′), 290/267 (5′-CCA GTC GCC TCC CCT CGC CTC TTG-3′) and 368/347 (5′-TCC CCC GCT TAA TAC TGA CGC T-3′) were used as antisense PCR primers. DNA products were separated on an agarose gel and extracted with the QIAEX II DNA isolation system according to the manufacturer's instructions (Qiagen). Transcription reactions were performed with the Megashortscript T7 transcription kit (Ambion) and transcripts were purified over a G50 spin column. Cy5 or Cy3 labeling was performed post-transcriptionally using the ULYSIS Cy5 or Cy3 nucleic acid labeling kit (Kreatech). RNA concentration and labeling density (approximately one Cy5 or Cy3 group per 34 nucleotides) were determined by spectrophotometry.

### Array synthesis

Arrays comprising every possible 5–25mer antisense oligonucleotide scanning HIV sequences +1 to +331 were synthesized using solid phase DNA chemistry. Oligonucleotides were synthesized *in situ* by delivery of phosphoramidite precursors to the surface of a derivatized

microscope glass slide using inkjet technology. Standard DNA chemistry couples the precursors to the array surface of the growing DNA chain and repeated rounds of synthesis create oligonucleotides of specific sequence and length that are spatially addressable (21,22).

### Array hybridization, data acquisition and analysis

The differential capacity of two different transcripts, one labeled with Cy5 and the other with Cy3, for antisense oligonucleotide binding was measured in a single array hybridization reaction. A control hybridization reaction was performed in parallel in which the fluorophore labels of the transcripts were switched. Hybridizations were performed in 280 μl reactions containing TEN buffer (100 mM NaCl, 1 mM EDTA, 10 mM Tris pH 8.0), supplemented with 1% Triton X-100. Transcripts (12 fmol RNA or 100 fmol Cy5 or Cy3) were heated to 85°C and slowly cooled to room temperature. Array hybridizations were performed for 16 h at room temperature under constant rotation of the hybridization chamber. Hybridization chambers were emptied and disassembled and the array was quickly placed in a 50 ml Falcon tube containing 1× TEN buffer with 0.1% Triton, washed for 30 s by inverting the tube, placed in a second tube containing 0.1× TEN and washed for a further 30 s with mixing. The array was slowly pulled from the final washing solution while being blown dry with a nitrogen gun at close range. The arrays were subsequently scanned in a confocal array scanner (Agilent) at 10 μm resolution with feature extraction software (Agilent) at default settings. The Agilent feature extraction software package was used to automatically detect oligonucleotide features in the hybridization image. The fluorescence intensities (arbitrary units) were exported to a tab-delimited text file and were imported in spreadsheet software for detailed analysis. The fluorescence intensities were corrected by subtraction of the local background measured around each feature. Features were rejected if the hybridization signal is not well above background. We only used background-corrected signals that are greater than 2.6 times the standard deviation of the background signal, which is consistent with a 99% confidence interval. The linear range of the measurements is approximately between 20 and 45 000 fluorescence units. The uniformity of the pixels within a feature signal was analyzed and features were rejected if either the pixels within a feature signal or the local background signal were not uniform (Agilent software, standard settings). Hybridization signals that did not pass these stringent quality tests such as uniformity of hybridization signal or low background signal, or that were at or below background level were discarded. To compare the hybridization patterns between transcripts, the fluorescence intensities were divided by the median fluorescence intensity to correct for differences in the Cy3 and Cy5 label and differences in labeling efficiency. All data on the oligonucleotide array experiments and subsequent analysis are available on our website (http://www.berkhoutlab.com/ooms2003array.xls).

### Gel mobility-shift assay

Oligonucleotide 58–79 (5′-CTT TAT TGA GGC TTA AGC AGT G-3′) was 5′-end labeled with the kinaseMax kit (Ambion) in the presence of 1 μl of [γ-$^{32}$P]ATP (0.37 MBq/μl, Amersham Biosciences). Transcriptions were made with the Megashortscript T7 transcription kit (Ambion) ethanol precipitated. The transcript concentration was determined by spectrophotometry. Transcripts were diluted in TEN buffer (100 mM NaCl, 1 mM EDTA, 10 mM Tris pH 8.0) to 0, 0.05, 0.25, 0.50 and 2.5 μM concentrations. Renaturation of the transcripts was performed in 18 μl samples by heating to 85°C and slowly cooling to room temperature. The 5′-labeled oligonucleotide 58–79 was added (5 nM) and the sample was incubated for 2 h at room temperature. After adding 4 μl of non-denaturing loading buffer (30% glycerol with bromophenol blue), samples were analyzed on a native 10% acrylamide gel (19:1 Serva) in 0.25× TBE. Electrophoresis was performed at 200 V at room temperature and the gel was subsequently dried. Quantification of the percentage shifted oligonucleotide was performed on a Phosphor Imager (Molecular Dynamics). The percentage shifted DNA oligonucleotide was calculated as the amount of shifted DNA oligonucleotide divided by the amount of shifted oligonucleotide + free oligonucleotide.

## RESULTS

### Microarray-based structure probing of HIV-1 RNA

The initial description of the alternative folding of HIV-1 leader RNA was based on a profound difference in the mobility of transcripts 1–245 and 1–290 on non-denaturing gels (7). In fact, the longer 1–290 transcript migrates faster than the 1–245 transcript, which is due to the formation of the stable LDI conformation when the downstream DIS sequences are included in the transcript (4). The poly(A) domain is directly involved in the long-distance base pairing that makes the LDI, but this domain folds into a hairpin in the BMH conformation. We therefore used these BMH (1–245) and LDI (1–290) transcripts (Fig. 1C) to probe for differences in accessibility to antisense oligonucleotides. These two transcripts, one labeled with Cy3 and the other with Cy5, were analyzed in a single hybridization reaction (Fig. 2). The binding results are shown in Figure 3A. There are major differences in oligonucleotide hybridization to the transcripts, e.g. the stable TAR hairpin is not very accessible, but the PBS and DIS regions are highly accessible for DNA oligonucleotides.

### Detection of alternative HIV-1 RNA conformations

We specifically wanted to use the array method to probe for differential oligonucleotide binding to identical leader RNA sequences that differentially fold the LDI or BMH conformation. We therefore calculated the ratio of oligonucleotide binding to the two RNA transcripts [ratio LDI (1–290):BMH (1–245)] from the crude hybridization yields (Fig. 3B). Indeed, a discrete segment within the poly(A) domain shows differential accessibility for the antisense DNA. Specifically, the antisense 55–79 DNA probe binds with at least 7-fold greater efficiency to LDI (1–290) than to BMH (1–245). Flanking oligonucleotides show a similar effect, but the differential accessibility quickly fades out in the neighboring upstream and downstream sequences. A smaller hybridization difference between the two transcripts is observed for the TAR sequence. The oligonucleotide complementary to HIV-1 leader positions 7–31 shows a 3-fold better hybridization to
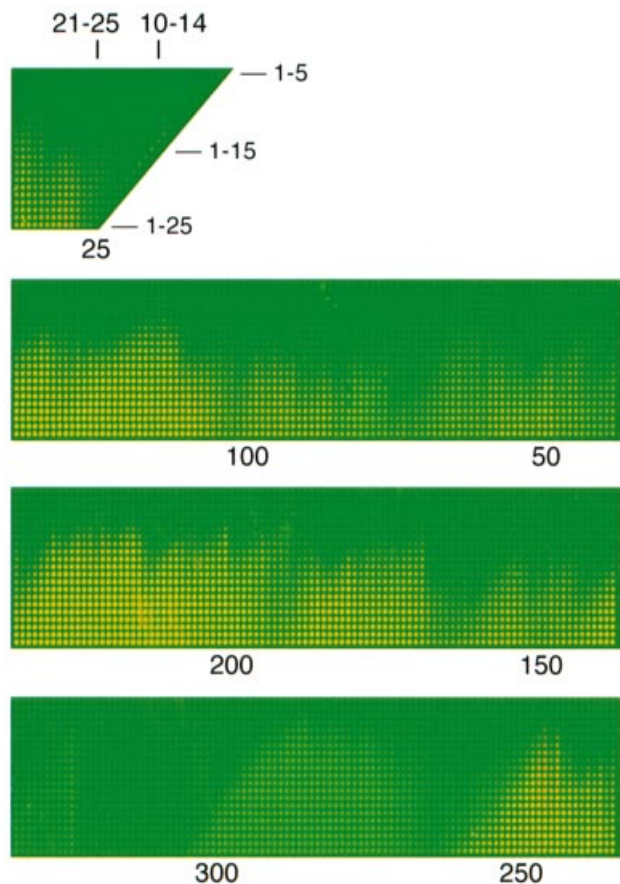
**Figure 2.** Scan of the oligonucleotide array with fluorescently labeled transcripts LDI 1–290 (green) and BMH 1–245 (red). The array contains all possible 5–25 nt oligonucleotides complementary to the HIV-1 leader RNA. The first row contains all 5mers, starting with oligonucleotide 1–5 on the upper right. The 5′ end positions of the oligonucleotides and the specific target positions on the HIV-1 target sequence are indicated.



**Figure 3.** (**A**) Hybridization pattern of the LDI 1–290 and BMH 1–245 transcripts. The hybridization yields (fluorescence intensities in arbitrary units) of 25 nt antisense probes are plotted against the start site of the HIV-1 target sequence. (**B**) Differential oligonucleotide hybridization to HIV-1 RNA in the LDI versus BMH conformation. The median-corrected differential hybridization yield of 1–290 LDI versus 1–245 BMH is plotted against the start site of the HIV-1 target sequence. The 1 value means that there is no difference in hybridization between the two transcripts. Values other than 1 indicate preferential binding to either the LDI or BMH transcript as indicated on the *y*-axis.

the BMH than to the LDI structure. No other difference in accessibility of the two transcripts is observed, except for a small increase in hybridization to the 3′ end of the 1–245 BMH transcript, which is probably caused by the inability of this region to adopt a stable RNA structure in the absence of downstream sequences.

The prominent difference in hybridization of the poly(A) domain between the two transcripts is solely caused by inclusion of the short downstream region 246–290. This finding is consistent with the structure models because the poly(A) region is base-paired with the DIS sequence in the LDI structure (Fig. 1B). The poly(A) domain folds into an RNA hairpin as part of the BMH structure in the absence of the DIS sequences. Inspection of the RNA secondary structure models in the LDI and BMH fold readily provides an explanation for the observed difference in hybridization of the 55–79 oligonucleotide. The RNA target contains 14 unpaired nucleotides in the LDI structure and only five in the BMH structure (Fig. 4).

Because downstream leader sequences were recently shown to contribute to both LDI and BMH foldings (6), we wanted to confirm the probing results obtained for the LDI–BMH switch
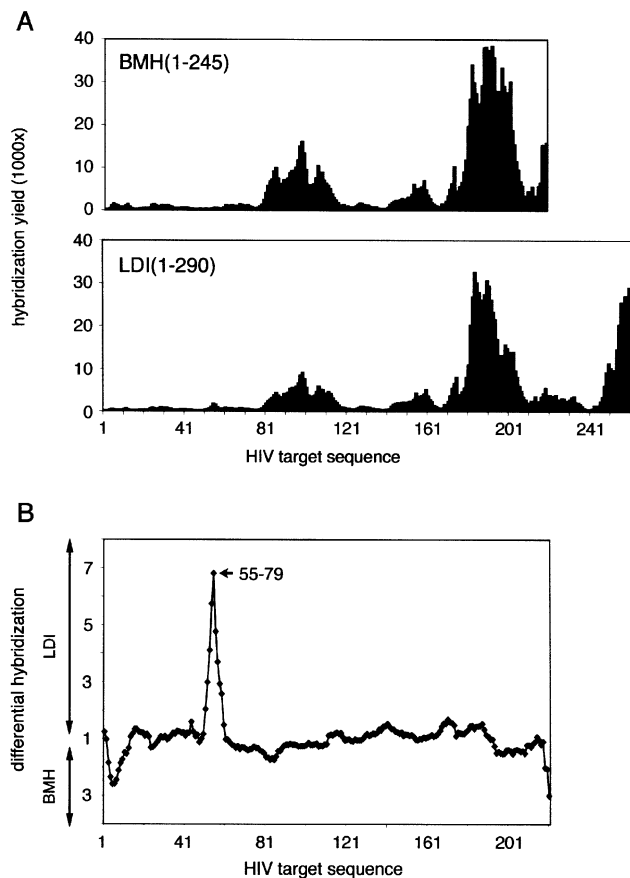
in the context of the full-length leader RNA. For this, we synthesized the wild-type 1–368 transcript that adopts the ground state LDI conformation, and stabilized the poly(A) hairpin as in mutant A to drive this transcript into the BMH form (see Fig. 4). Previous studies demonstrated that this A mutant transcript stably adopts the BMH conformation (4). The array hybridization results up to leader position 307 are presented for the 25mer oligonucleotides in Figure 5A. We excluded the data in the 67–100 region for the A mutant because the mutations destroy complementarity to the oligonucleotides of the array in this region (indicated by an open bar). The results are very similar to the experiment with the shorter RNA templates (Fig. 3A), showing the inaccessibility of the TAR domain and high accessibility of the PBS and the DIS domains.

Differential oligonucleotide accessibility towards the LDI is apparent for these transcripts in the poly(A) region (Fig. 5B). The results with the full-length leader RNA are strikingly similar to those obtained with the shorter transcripts, and the
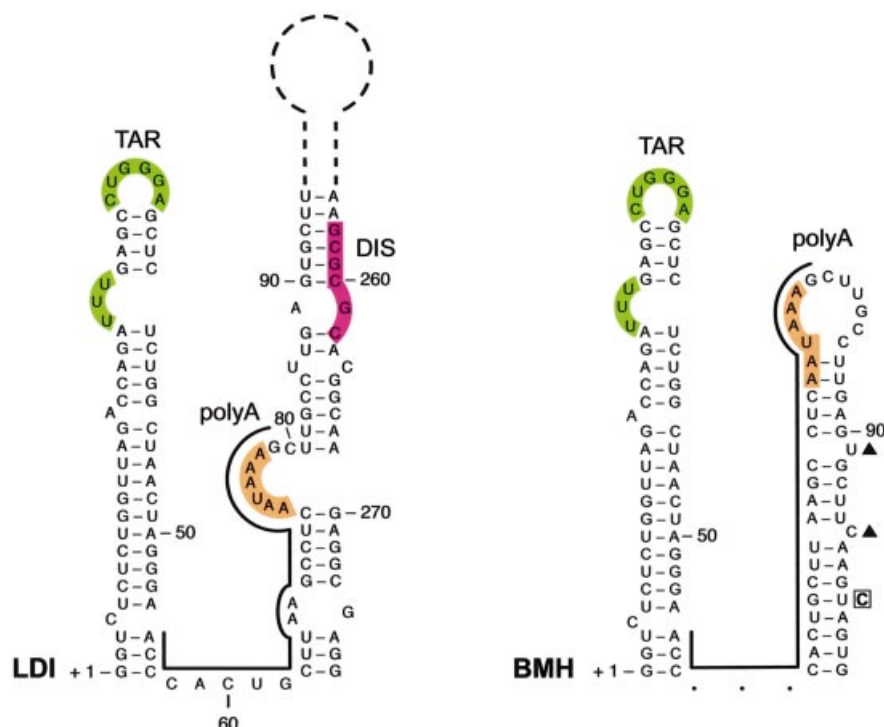
**Figure 4.** Secondary structure of the poly(A) region in the LDI and BMH conformations. Shown are the adjacent TAR and poly(A) domains. In the LDI structure, the poly(A) sequence is base-paired with the downstream DIS segment. In the BMH, poly(A) sequences form a distinct hairpin structure. A black line indicates the target sequence of the 55–79 oligonucleotide that binds preferentially to LDI. The target contains 14 unpaired nucleotides in the LDI and only five unpaired nucleotides in the BMH. The poly(A) mutant A analyzed in the 1–368 context (Fig. 5) has a deletion of residues U91 and C96, and the U100C substitution that result in significant stabilization of the poly(A) hairpin.

most dramatic effect is again observed for the oligonucleotide 55–79. The only difference seems to be the magnitude of this differential binding, which is >72-fold in this experiment compared with 7-fold in the previous experiment. It is likely that stabilization of the poly(A) hairpin in mutant A has a direct negative impact on binding of the antisense DNA in the BMH conformation, thus emphasizing the difference from the LDI conformation. In addition, the stability of the LDI conformation may be higher in the 1–368 transcript compared with the 1–290 transcript (4). Other findings are also very similar for the short and long HIV-1 leader transcripts. For instance, preferential binding to the TAR element in the BMH context is observed in both experimental systems. The region with the highest differential binding towards the BMH is located in the DIS domain. This 2-fold increase in hybridization of oligonucleotide 252–276 to the BMH transcript, resulting from stabilization of the poly(A) hairpin, is consistent with the LDI and BMH structure models. Indeed, the RNA segment that is targeted by the 252–276 oligonucleotide has only eight unpaired nucleotides in the LDI structure model, but 12 unpaired nucleotides in the BMH conformation (Fig. 6).

### Differential oligonucleotide binding to LDI and BMH in a gel mobility-shift assay

To determine if the differential binding of oligonucleotide 55–79 can also be demonstrated by gel mobility-shift assays, we performed this assay with ^{32}P-labeled oligonucleotide 58–79 (5 nM) and different concentrations of LDI and BMH

leader transcripts (0, 0.05, 0.25, 0.50 and 2.5 μM) (Fig. 7A). We used the 58–79 instead of the 55–79 oligonucleotide because the 3′ end, which is near to the array surface, is not involved in hybridization. This is probably due to steric hindrance. The gel mobility-shift experiment confirms that oligonucleotide 58–79 has a higher affinity for the 1–290 LDI than the 1–245 BMH transcript (Fig. 7A and B). The results with the BMH and LDI transcripts in the 1–368 context are also very similar to the microarray data. The A mutant BMH transcript is unable to shift the DNA, even at the highest RNA concentration tested, but the wild-type LDI transcript shifts the DNA efficiently. The absolute oligonucleotide binding defect of the A mutant is probably caused by more stable base-pairing of part of its target sequence in the poly(A) hairpin. In general, these results demonstrate that the binding data obtained with microarrays do not differ from those from more traditional binding assays.

### Microarray hybridization is consistent with classical RNA structure probing

A positive correlation between the activity of single strand-specific RNases and accessibility for antisense DNA binding has been described previously (23). The requirement for unpaired regions for efficient DNA oligonucleotide hybridization has also been observed (16,24). The fact that the HIV-1 leader RNA structure has been intensively studied provides an opportunity to examine the role of the secondary structure of the target RNA in the efficiency of DNA oligonucleotide
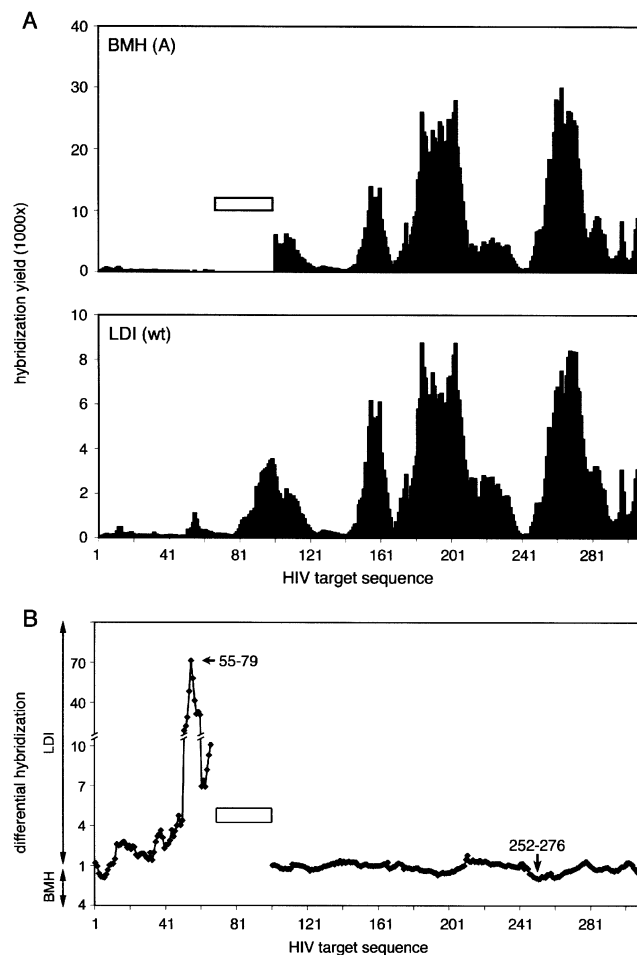
**Figure 5.** (**A**) Hybridization patterns of LDI 1–368 wt and BMH 1–368 A transcripts. The hybridization yields (fluorescence intensities in arbitrary units) of 25mer oligonucleotides are plotted against the start site of the HIV-1 target sequence. Because the 1–368 A transcript contains mutations and a deletion in the poly(A) sequence, no data are shown for this region (indicated by the open bar). (**B**) Differential hybridization of oligonucleotides to transcripts 1–368 wt (LDI) and 1–368 A (BMH). The median-corrected differential hybridization yields of 1–368 wt and 1–368 A are plotted against the start site of 25mer nucleotides on the HIV-1 target sequence. The 1 value means that there is no difference in hybridization between the two transcripts. Values other than 1 indicate preferential binding to either the LDI or BMH transcript as indicated on the y-axis. The open bar indicates the positions where the oligonucleotides have mismatches to the mutant A transcript.



**Figure 6.** Secondary structure of the HIV-1 DIS domain in the LDI and BMH conformation. The LDI structure of the 1–368 transcript differs in some details from that of the 1–290 transcript shown in Figure 4 (4,6). In the LDI structure, the DIS sequence is base-paired with the upstream poly(A) segment. In the BMH, it forms a distinct hairpin structure. A black line indicates the target sequence of the 252–276 oligonucleotide that preferentially binds the BMH. The target sequence contains eight unpaired nucleotides in the LDI and 12 unpaired nucleotides in the BMH.

hybridization (6,25,26). Structure probing analysis on the HIV-1 MAL isolate was not included because this isolate contains a 20 nt insertion 3′ of the PBS and additional sequence differences compared with the HIV-1 LAI isolate used in the oligonucleotide arrays (25).We recently reported a detailed structure probing analysis of both the LDI and the BMH structures of the full-length HIV-1 LAI leader transcript (6). The structure probing was performed with dimethylsulfate (DMS; specific for unpaired adenines and cytosines) and kethoxal (unpaired guanines). To compare this data set with the array data, we plotted the number of modified nucleotides (DMS + kethoxal) for each 25 nt segment against the array hybridization values of the wild-type full-length leader RNA
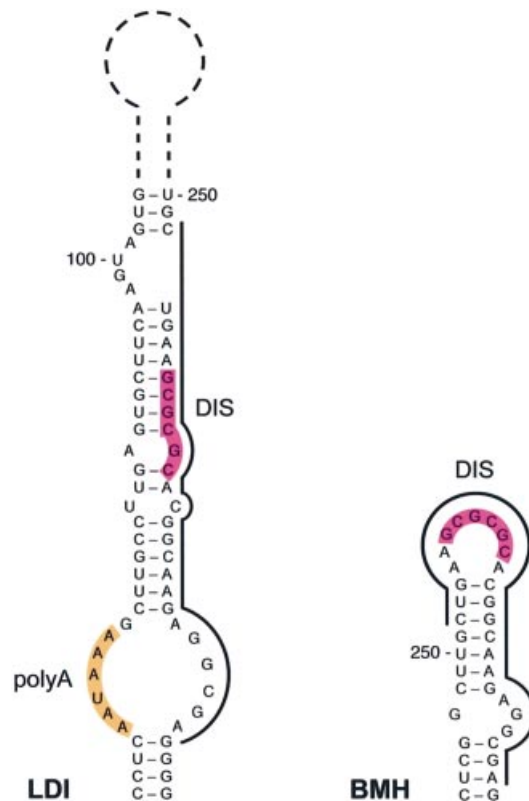
(Fig. 8A). This comparison reveals a strong positive correlation between the number of hits with single strand-specific chemicals and hybridization to antisense DNA oligonucleotides (Spearman's rank correlation coefficient 0.705; $P < 0.01$).

Another detailed probing study was performed with the RNases T1 and T2, which preferentially cleave at unpaired guanines and adenines, respectively, and the cobra venom V1 enzyme that prefers stacked nucleotides and/or nucleotides involved in Watson–Crick base-pairing (26). We plotted the number of T1 and T2 cleavages against the array data (Fig. 8B), and a positive correlation is again apparent, although less prominent than for the chemical probing (Spearman's rank correlation coefficient 0.35; $P < 0.01$). This may be due to the bulkiness of the enzyme probes, which may be unable to interact with all the unstructured nucleotides. Furthermore, a negative correlation is apparent for the V1 reactivity and oligonucleotide binding (Fig. 8C, Spearman's rank correlation coefficient –0.354; $P < 0.01$). The oligonucleotide hybridization defect is absolute for target RNA segments with at least eight V1 hits. We conclude that the oligonucleotide hybridization efficiency is strongly influenced by the secondary structure of the target RNA.
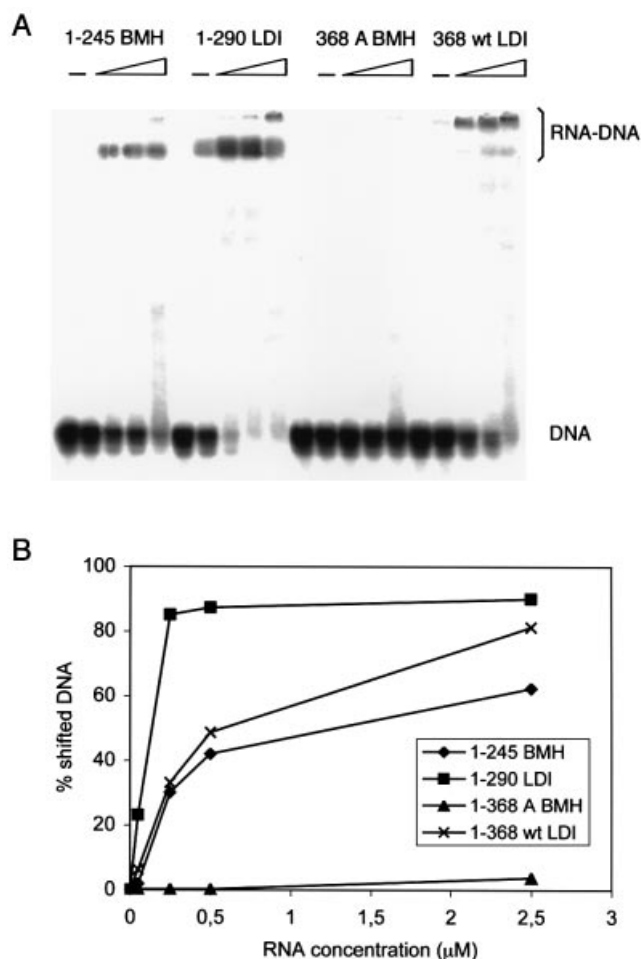
**Figure 7.** Gel mobility-shift assay. (**A**) ^{32}P-labeled oligonucleotide 58–79 (5 nM) was incubated with different concentrations of leader transcripts (0, 0.05, 0.25, 0.50 and 2.5 µM). The samples were analyzed on a native 10% acrylamide gel in 0.25× TBE. The positions of the free DNA oligonucleotide and the RNA–DNA duplexes are indicated. (**B**) Quantification of the percentage shifted DNA oligonucleotide, which was calculated by shifted DNA oligonucleotide/(shifted oligonucleotide + free oligonucleotide).

## DISCUSSION

The use of scanning arrays of antisense oligonucleotides allows the systematic analysis of target RNA accessibility with all possible oligonucleotides (1–25, 2–26, 3–27, …) and many length variations of the DNA probes. This novel technique provides an empirical approach for the selection of potent antisense oligonucleotides (12,13,27–29), and it can also identify efficient targets for RNA interference approaches (30,31). We demonstrate in this study that the results of microarray binding experiments are comparable with more traditional binding experiments such as gel mobility-shift assays. Furthermore, there is recent evidence that the *in vitro* selected antisense molecules are also efficient antisense reagents *in vivo*, despite the complexity of the intracellular environment (14). The array technology is ideally suited to perform a systematic analysis of the influence of the target RNA structure on the binding of oligonucleotides (15,16). We probed the untranslated leader of the HIV-1 genome for which
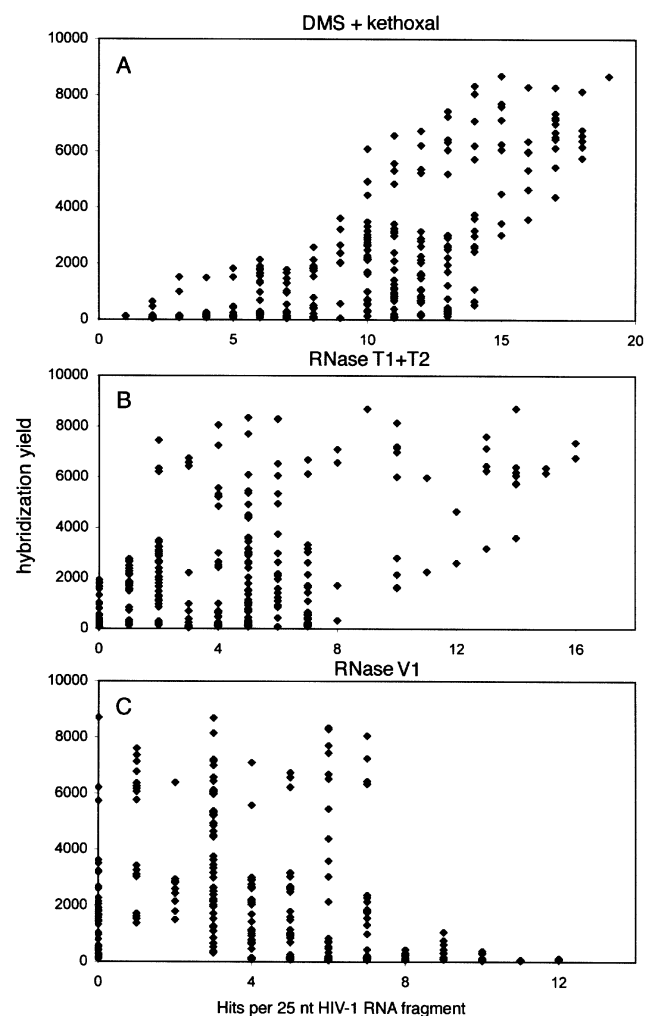


**Figure 8.** Antisense oligonucleotide binding versus RNA structure probing. (**A**) Correlation between oligonucleotide annealing and unpaired nucleotides in the target sequence, as determined by the number of nucleotides within the 25mer target sequence cleaved by the single strand-specific chemicals kethoxal and DMS (31) (Spearman's rank correlation coefficient 0.705; *P* < 0.01). (**B**) Correlation between oligonucleotide annealing and unpaired nucleotides in the target sequence, as determined by the number of nucleotides within the 25mer target sequence cleaved by the single strand-specific enzymes RNase T1 and T2 (25) (Spearman's rank correlation coefficient 0.35; *P* < 0.01). (**C**) Correlation between oligonucleotide annealing and paired nucleotides in the target sequence, as determined by the number of nucleotides within the 25mer target sequence cleaved by the double strand-specific enzyme cobra venom V1 (25) (Spearman's rank correlation coefficient –0.354; *P* < 0.01).

a wealth of experimental data is already available, and the results indicate that arrays can indeed provide global information on the RNA structure. For instance, the fairly stable TAR hairpin is inaccessible for antisense DNA, and the largely unstructured PBS domain is highly accessible. However, this approach will probably not yield the detailed information on the RNA secondary structure obtained by methods such as enzymatic and chemical structure probing.

We also tested whether the array method is suited for identification of structural differences in transcripts that adopt alternative conformations. We previously demonstrated such an RNA switch for the untranslated HIV-1 leader RNA (4). In

this study, we confirmed that the LDI–BMH switch in HIV-1 RNA directly affects the folding of the poly(A) and DIS domains. For this, we compared the wild-type LDI transcript with two mutants that are forced into the alternative BMH folding: a DIS-minus transcript (1–245) that cannot form the LDI, and the A mutant that stabilizes the poly(A) hairpin of the BMH conformation. Deleting the DIS sequence changes the accessibility of the poly(A) hairpin, and stabilization of the poly(A) hairpin changes the accessibility of the DIS domain. Comparing classical enzymatic and chemical structure probing with the novel array data demonstrates the strong influence of the target RNA structure on oligonucleotide hybridization. Thus, differences in hybridization patterns directly reflect RNA structural changes, and the results are fully consistent with the current structure models. The poly(A) and the DIS regions interact in the LDI conformation, but these regions form distinct hairpins in the BMH structure. We observed an additional difference in upstream TAR sequences. Although the TAR hairpin is folded identically in the LDI and BMH structure models (Fig. 1B), we previously demonstrated that correct TAR folding has an influence on the LDI–BMH equilibrium, possibly via tertiary interactions (8). The results of the array analysis are consistent with this idea, although they do not provide further insight into the molecular details of this TAR effect.

The antisense oligonucleotides we identified are of interest for several reasons. They include the most efficient binders, e.g. oligonucleotides in the PBS and DIS regions, but also oligonucleotides that bind preferentially to one of the alternative leader RNA conformations; e.g. the LDI-specific binder (55–79) that targets the poly(A) domain and the preferential BMH binder (252–276) in the DIS domain. If these molecules affect the natural LDI–BMH equilibrium within the virus-infected cell, we expect that the coordination of the viral gene expression program will be disturbed. In general, such differentially binding oligonucleotides could be used as tools to dissect the role of alternative RNA conformations *in vivo*.

## REFERENCES

1. Berkhout,B. (1996) Structure and function of the human immunodeficiency virus leader RNA. *Prog. Nucleic Acid Res. Mol. Biol.*, **54**, 1–34.
2. Berkhout,B. (2000) Multiple biological roles associated with the repeat (R) region of the HIV-1 RNA genome. *Adv. Pharmacol.*, **48**, 29–73.
3. Beerens,N. and Berkhout,B. (2002) Strict regulation of HIV-1 reverse transcription. *Curr. Top. Virol.*, **2**, 115–127.
4. Huthoff,H. and Berkhout,B. (2001) Two alternating structures for the HIV-1 leader RNA. *RNA*, **7**, 143–157.
5. Berkhout,B., Ooms,M., Beerens,N., Huthoff,H., Southern,E. and Verhoef,K. (2002) *In vitro* evidence that the untranslated leader of the HIV-1 genome is an RNA checkpoint that regulates multiple functions through conformational changes. *J. Biol. Chem.*, **277**, 19967–19975.
6. Abbink,T.E. and Berkhout,B. (2003) A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. *J. Biol. Chem.*, **278**, 11601–11611.
7. Berkhout,B. and van Wamel,J.L.B. (2000) The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. *RNA*, **6**, 282–295.
8. Huthoff,H. and Berkhout,B. (2001) Mutations in the TAR hairpin affect the equilibrium between alternative conformations of the HIV-1 leader RNA. *Nucleic Acids Res.*, **29**, 2594–2600.
9. Lanchy,J.M., Rentz,C.A., Ivanovitch,J.D. and Lodmell,J.S. (2003) Elements located upstream and downstream of the major splice donor site influence the ability of HIV-2 leader RNA to dimerize *in vitro*. *Biochemistry*, **42**, 2634–2642.
10. Dirac,A.M., Huthoff,H., Kjems,J. and Berkhout,B. (2002) Regulated HIV-2 RNA dimerization by means of alternative RNA conformations. *Nucleic Acids Res.*, **30**, 2647–2655.
11. Huthoff,H. and Berkhout,B. (2002) Multiple secondary structure rearrangements during HIV-1 RNA dimerization. *Biochemistry*, **41**, 10439–10445.
12. Milner,N., Mir,K.U. and Southern,E.M. (1997) Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat. Biotechnol.*, **15**, 537–541.
13. Sohail,M. and Southern,E.M. (2000) Antisense arrays. *Mol. Cell Biol. Res. Commun.*, **3**, 67–72.
14. Sohail,M., Hochegger,H., Klotzbucher,A., Guellec,R.L., Hunt,T. and Southern,E.M. (2001) Antisense oligonucleotides selected by hybridisation to scanning arrays are effective reagents *in vivo*. *Nucleic Acids Res.*, **29**, 2041–2051.
15. Sohail,M., Akhtar,S. and Southern,E.M. (1999) The folding of large RNAs studied by hybridization to arrays of complementary oligonucleotides. *RNA*, **5**, 646–655.
16. Mir,K.U. and Southern,E.M. (1999) Determining the influence of structure on hybridization using oligonucleotide arrays. *Nat. Biotechnol.*, **17**, 788–792.
17. Das,A.T., Klaver,B., Klasens,B.I.F., van Wamel,J.L.B. and Berkhout,B. (1997) A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. *J. Virol.*, **71**, 2346–2356.
18. Peden,K.W.C. and Martin,M.A. (1996) In Karn,J. (ed.), *HIV: A Practical Approach*. IRL Press at Oxford University Press, Oxford, UK, Vol. 1, pp. 21–45.
19. Klasens,B.I.F., Das,A.T. and Berkhout,B. (1998) Inhibition of polyadenylation by stable RNA secondary structure. *Nucleic Acids Res.*, **26**, 1870–1876.
20. Das,A.T., Klaver,B. and Berkhout,B. (1999) A hairpin structure in the R region of the human immunodeficiency virus type 1 RNA genome is instrumental in polyadenylation site selection. *J. Virol.*, **73**, 81–91.
21. Hughes,T.R., Mao,M., Jones,A.R., Burchard,J., Marton,M.J., Shannon,K.W., Lefkowitz,S.M., Ziman,M., Schelter,J.M., Meyer,M.R. *et al.* (2001) Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.*, **19**, 342–347.
22. Shoemaker,D.D., Schadt,E.E., Armour,C.D., He,Y.D., Garrett-Engele,P., McDonagh,P.D., Loerch,P.M., Leonardson,A., Lum,P.Y., Cavet,G. *et al.* (2001) Experimental annotation of the human genome using microarray technology. *Nature*, **409**, 922–927.
23. Matveeva,O., Felden,B., Audlin,S., Gesteland,R.F. and Atkins,J.F. (1997) A rapid *in vitro* method for obtaining RNA accessibility patterns for complementary cDNA probes: correlation with an intracellular pattern and known RNA structures. *Nucleic Acids Res.*, **25**, 5010–5016.
24. Lima,W.F., Monia,B.P., Ecker,D.J. and Freier,S.M. (1992) Implication of RNA structure on antisense oligonucleotide hybridization kinetics. *Biochemistry*, **31**, 12055–12061.
25. Baudin,F., Marquet,R., Isel,C., Darlix,J.L., Ehresmann,B. and Ehresmann,C. (1993) Functional sites in the 5′ region of human immunodeficiency virus type 1 RNA form defined structural domains. *J. Mol. Biol.*, **229**, 382–397.
26. Damgaard,C.K., Dyhr-Mikkelsen,H. and Kjems,J. (1998) Mapping the RNA binding sites for human immunodeficiency virus type-1 Gag and NC proteins within the complete HIV-1 and -2 untranslated leader regions. *Nucleic Acids Res.*, **26**, 3667–3676.

27. Southern,E.M., Case-Green,S.C., Elder,J.K., Johnson,M., Mir,K.U., Wang,L. and Williams,J.C. (1994) Arrays of complementary oligonucleotides for analysing the hybridisation behaviour of nucleic acids. *Nucleic Acids Res.*, **22**, 1368–1373.
28. Milner,N., Mir,K.U. and Southern,E.M. (1997) Selecting effective antisense reagents on combinatorial oligonucleotide arrays. *Nat. Biotechnol.*, **15**, 537–541.
29. Southern,E.M., Milner,N. and Mir,K.U. (1997) Discovering antisense reagents by hybridization of RNA to oligonucleotide arrays. *Ciba Found. Symp.*, **209**, 38–46.

30. Bohula,E.A., Salisbury,A.J., Sohail,M., Playford,M.P., Riedemann,J., Southern,E.M. and Macaulay,V.M. (2003) The efficacy of small interfering RNAs targeted to the type 1 IGF receptor is influenced by secondary structure in the IGF1R transcript. *J. Biol. Chem.*, **278**, 15991–15997.
31. Vickers,T.A., Koo,S., Bennett,C.F., Crooke,S.T., Dean,N.M. and Baker,B.F. (2003) Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *J. Biol. Chem.*, **278**, 7108–7118.