A Microbead-based System for Identifying and Characterizing RNA-Protein Interactions by Flow Cytometry*

Alexander S. Brodsky‡§ and Pamela A. Silver

We present a high throughput, versatile approach to identify RNA-protein interactions and to determine nucleotides important for specific protein binding. In this approach, oligonucleotides are coupled to microbeads and hybridized to RNA-protein complexes. The presence or absence of RNA and/or protein fluorescence indicates the formation of an oligo-RNA-protein complex on each bead. The observed fluorescence is specific for both the hybridization and the RNA-protein interaction. We find that the method can discriminate noncomplementary and mismatch sequences. The observed fluorescence reflects the affinity and specificity of the RNA-protein interaction. In addition, the fluorescence patterns footprint the protein recognition site to determine nucleotides important for protein binding. The system was developed with the human protein U1A binding to RNAs derived from U1 snRNA but can also detect RNA-protein interactions in total RNA backgrounds. We propose that this strategy, in combination with emerging coded bead systems, can identify RNAs and RNA sequences important for interacting with RNA-binding proteins on genomic scales. ***Molecular & Cellular Proteomics 1:922–929, 2002.***

RNA-protein interactions are a central component of post-transcriptional regulation at multiple levels including RNA processing, transport, and translation. The sequenced human genome reveals hundreds of potential RNA-binding proteins (1). A critical step toward understanding the function of RNA-binding proteins is to identify and determine how they interact with their target RNAs.

Several strategies have been developed to identify RNA-protein interactions. Genetic approaches include three-hybrid screens (2), phage display (3), and TRAP (translational repression assay procedure) (4) to identify proteins that bind a specific RNA. However, these strategies are generally not applicable to larger RNAs and not suitable for determining binding constants. SELEX (systematic evolution of ligands by exponential enrichment) can identify high affinity RNA sequences that may or may not reflect the biologically relevant recognition site (5).

Recently immunoprecipitation has been combined with microarray analysis to identify candidate RNAs bound to proteins (6, 7). This approach is very promising for inspection of RNA-protein interactions on a genome-wide scale. However, it relies on the ability to preserve stable interactions during immunoprecipitation; many potentially weak interactions may be lost. In addition, extensive motif searching together with additional experimentation may be required to validate the biological significance of any interactors (8).

Recent advances in bead coding technologies to create high complexity platforms are leading to the development of new approaches for high throughput screening studies that could be amenable to the study of RNA-protein interactions (9–12). In principle, nucleic acid hybridization on microbeads offers a number of advantages over DNA chips including shorter hybridization times and better control of binding conditions (13). Therefore, we have developed a new equilibrium binding method on microbeads for elucidating the recognition site of an RNA-binding protein on its cognate RNA. The approach uses oligonucleotide-coupled microbeads and fluorescently labeled RNAs and proteins to monitor RNA-protein binding by flow cytometry. To develop the system for screening RNA-protein interactions, we demonstrate how this approach can be used to identify and characterize the interaction between the spliceosomal protein U1A and a hairpin derived from U1 snRNA as well as detect specific RNAs from total RNA populations.

EXPERIMENTAL PROCEDURES

Plasmids—The U1A test transcript was constructed by annealing overlapping oligos and ligating the annealed product into pDP19 (Ambion) to create plasmid pPS2702. The oligo sequences are: AATTCTTTATCTTCAAAGTTGTCTGTCCAAGATTTGGACTTGTCCGGAGTGCAATGGACG, AAGGACAAGCGTGTCTTCATCAGAGTTGACTTCACTCGAG, GACAAGTCCAAATCTTGGACAGACAACTTTGAAGATAAAG, and GATCCTCGAGTGAAGTCAACTCTGATGAAGACACGCTTGTCCTTCGTCCATTGCACTCCG.

U1A-green fluorescent protein (GFP)[1] was PCR-amplified from pPS2035 and ligated into prSETB (Invitrogen) to create pPS2699. The 96A→G U1A point mutant was constructed by using Stratagene's QuikChange system to create pPS2703. 77C→G was constructed by

From the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School and The Dana-Farber Cancer Institute, Boston, Massachusetts 02115

Received, October 1, 2002, and in revised form, November 27, 2002

Published, MCP Papers in Press, December 1, 2002, DOI 10.1074/mcp.T200010-MCP200

[1] The abbreviation used is: GFP, green fluorescent protein.

ligating annealed oligonucleotides into pDP19 as described above. All constructs were verified by sequencing.

Transcription and RNA Preparations—PPS2702 and pPS2703 were linearized with *Eco*RI and subsequently transcribed with Ambion's T3 polymerase Maxiscript kit. Labeling with ^{32}P verified a product of the expected size, and subsequent transcription reactions were purified by G-50 spin columns or phenol extractions followed by multiple ethanol precipitations. Texas Red-5-UTP (Molecular Probes) was incorporated during transcription and purified with G-50 spin columns. Total RNA from HeLa cells was prepared by the TRIzol method with high salt precipitations to reduce background GFP fluorescence. Yeast RNA was isolated by a hot phenol method. RNA concentrations were determined by UV spectrometry.

U1A-GFP Purification—Cells were grown to 0.5 optical density before induction with 1 mM isopropyl-1-thio-β-D-galactopyranoside for 4 h. Cells were resuspended in 20 mM HEPES, 10 mM KCl, 0.1% IGEPAL and lysed with lysozyme followed by sonication. After centrifugation, lysate was applied to nickel columns, washed extensively, and eluted with imidazole. Green fractions were pooled and dialyzed into 10 mM HEPES, pH 7.6, 10 mM KCl, 0.1% IGEPAL. To remove the histidine tag, 1.25 units/μl enterokinase were added and incubated for >48 h at 25 °C. Enterokinase was removed with EKaway resin (Invitrogen). U1A-GFP was dialyzed into storage buffer (20 mM HEPES, 20 mM KCl, 0.1% IGEPAL, 10% glycerol). Concentrations were determined by comparing U1A-GFP to bovine serum albumin on Coomassie gels and by the Bio-Rad protein assay. Protein stored at −80 °C bound RNA similarly to fresh preparations (data not shown).

Bead Preparation—Before coupling, Dynal 2.8-μm magnetic streptavidin beads (M-280) were vortexed and/or sonicated to reduce aggregation. Similar to Dynal's recommended protocols, oligonucleotides were attached to beads with 1 nmol of oligonucleotide/9 × 10^6 beads/30 μl. Incubations longer than 5 h were required to reach maximum oligonucleotide density (data not shown). Similar procedures were used for the Spherotech 5.7-μm magnetic streptavidin beads. Oligonucleotides were synthesized with a 12-carbon spacer and 5′ biotin from two different sources: Dana-Farber Cancer Institute Molecular Biology Core Facilities and Integrated DNA Technologies, Inc. Oligonucleotides from each source behaved similarly.

Bead Binding Assays—Binding was performed in 20 mM HEPES, pH 7.5, 300 mM KCl, 0.1% IGEPAL, 10 ng/μl tRNA, 0.04 units/μl superase-IN (Ambion), and 20 ng/μl bovine serum albumin unless otherwise indicated. RNA was heated to 95 °C for 1 min and cooled on ice before being mixed with U1A-GFP for at least 20 min at room temperature before addition of 1 × 10^5 oligonucleotide-coupled beads. Binding reactions were incubated at 35 °C for at least 14 h with constant rotation unless otherwise indicated. Shorter incubations (<6 h) gave lower fluorescence.

Flow Cytometry and Data Analysis—A BD Biosciences Vantage was used to sort beads with both GFP and Texas Red signals. A FACScan was used to monitor GFP alone. Typically 5,000–10,000 beads were counted, and the peak channel, which indicates the maximum height of the bead population, is used to estimate the fluorescence intensity. To determine the percentage of the population shifted to higher fluorescence, cut-offs were set relative to the background fluorescence. Beads with fluorescence above the cut-off are counted in the shifted population. For the binding curves, pro Fit (Quantum Soft) was used to fit the fluorescence intensities to a Langmuir isotherm.

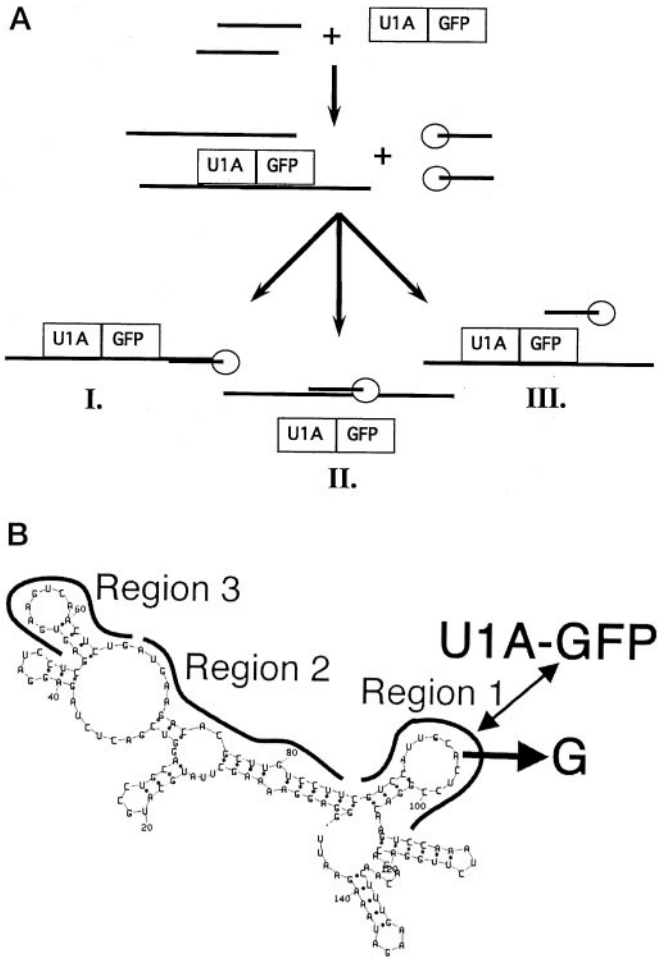

FIG. 1. **Experimental design.** *A*, schematic of the experiment. A U1A-GFP-RNA complex is formed and subsequently challenged with oligonucleotide beads. After reaching equilibrium, RNA and protein fluorescence on each microbead are determined by flow cytometry. The experiment can be performed with or without fluorescently labeled RNA. Three scenarios are possible. *I*, GFP signal is observed indicating the bead is coupled to an oligo complementary to the RNA target but distant from the U1A stem loop recognition site. *II*, no GFP signal is observed, but the oligo is hybridizing to the RNA. With labeled RNA, the RNA-oligo hybridization is detected. These oligos may be complementary to the U1A binding site. *III*, beads with neither GFP nor RNA fluorescence are observed, suggesting that these oligos do not hybridize to the RNA. These sequences may be non-complementary to the RNA. *B*, predicted secondary structure of U1A RNA constructed for these studies as determined by mFOLD (21). U1A-GFP binds to the hairpin derived from U1 snRNA as indicated. An A to G mutation (96A→G) reduces binding 1000-fold. *Region 1* is complementary to the binding site. Oligonucleotides complementary to other regions of the RNA, distant from the binding site, are also indicated. These sequences are predicted to hybridize to the RNA while U1A-GFP is binding, allowing the U1A RNA-protein interaction to be observed.

RESULTS

The experimental design for analysis of protein-RNA interactions uses oligonucleotides coupled to microbeads to probe RNA-protein interactions and is outlined in Fig. 1*A*. To carry out the analysis, a protein-RNA complex is first formed and then incubated with beads to which oligonucleotides complementary to the target RNA have been coupled. In the experiments described here, RNA is labeled with Texas Red, and the protein is a GFP fusion. After reaching equilibrium, the

TABLE I
Oligonucleotides used in this study
Mismatch nucleotides are in lowercase.

Name	Sequence
	Oligonucleotides complementary to U1A binding site
1.20[a]	TTGTCCGGAGTGCAATGGAC
1.17	GTCCGGAGTGCAATGGA
1.15	GTCCGGAGTGCAATG
	Oligonucleotides complementary to regions distant from U1A binding site
2.20	AAGGACAAGCGTGTCTTCAT
2.17	GGACAAGCGTGTCTTCA
2.15	GGACAAGCGTGTCTT
3.17	TCAGAGTTGACTTCACT
3.15	TCAGAGTTGACTTCA
	Mismatch oligonucleotides
2C.20[b]	AAGGACAAcCGTGTCTTCAT
2C.17	GGACAAcCGTGTCTTCA
2C.15	GGACAAcCGTGTCTT
3C.17	TCAGAcTTGACTTCACT
3C.15	TCAGAcTTGACTTCACT
	U1 snRNA oligonucleotides
4.17	CCCTGCCAGGTAAGTAT
4G.17	CCCTGCgAGGTAAGTAT

[a] The first number indicates the sequence that is being targeted, while the second number indicates the oligonucleotide length, *e.g.* 1.20 is an oligonucleotide complementary to region 1 with 20 nucleotides.

[b] The letter indicates a point mutation in the oligonucleotide disrupting hybridization to the RNA.

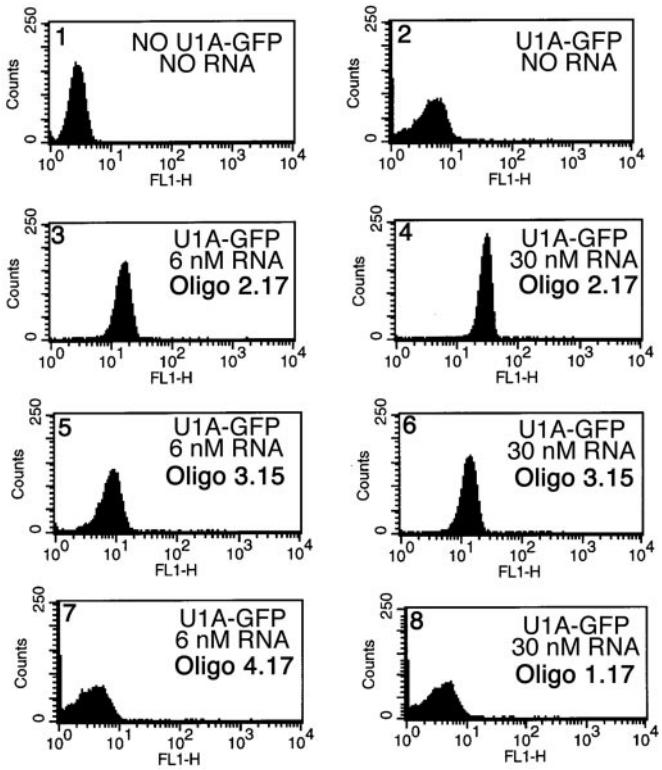

FIG. 2. **Fluorescence is RNA-dependent.** Histograms show the number of beads at different GFP fluorescence intensities. In the presence of RNA and U1A-GFP, the fluorescence intensity of the bead population increases, and a more homogeneous bead population is observed (compare *panels 2*, *3*, and *4*). Oligonucleotides targeting the binding site or not complementary to the RNA show nonspecific binding (compare *panels 2*, *7*, and *8*).

beads are analyzed in a flow cytometer for protein and RNA fluorescence. Beads are sorted into different categories as illustrated in Fig. 1*A*. 1) Beads with both GFP and RNA signals represent the RNA-protein interaction. The oligonucleotides on the beads hybridize to the RNA without interfering with protein binding. 2) RNA signal alone indicates RNA hybridization with no protein binding. The oligonucleotides on these beads may be competing with the protein to bind the same RNA sequences. 3) Some beads will have no detectable fluorescent signal. These oligonucleotide-coupled beads contain sequences that cannot hybridize to the RNA including those that are noncomplementary or contain point mutations. To quantitate the data, the mean fluorescence and/or the percentage of beads in the different categories is determined.

Development of the Bead Binding Assay—The system was developed with the human splicing protein U1A binding to the stem loop 2 derived from U1 snRNA. A U1A-GFP fusion protein including the first 94 amino acids of the RNA recognition motif was expressed with an N-terminal His_6 tag and a C-terminal GFP. The histidine tag was proteolytically cleaved to generate functional U1A-GFP. A 145-nucleotide RNA encoding random sequence and including the specific U1A hairpin was designed. The predicted secondary structure is shown in Fig. 1*B*. Gel shift mobility experiments confirmed that U1A-GFP binds the U1A RNA with a dissociation constant of ~35 nM in 150 mM KCl at 25 °C (data not shown), similar to that reported for the same 94-amino acid fragment in the absence of the GFP (14). Oligonucleotides complementary to different regions of the U1A RNA were designed as illustrated in Fig. 1*B* and listed in Table I. Oligonucleotides were synthesized with a 5′ biotin label and attached to streptavidin beads. Reproducible coupling conditions were devised to ensure similar oligonucleotide concentrations per bead as determined with ^{32}P-labeled oligos (data not shown). The oligonucleotide concentrations used in the binding assay are estimated to be between 1–2 nM. For all experiments, Dynal 2.8-μm diameter streptavidin beads were used unless otherwise indicated.

After the RNA-protein complex is formed, oligo-coupled beads are added. To reach equilibrium, incubations at 25 °C or 35 °C for longer than 6 h were necessary (data not shown) and typical incubations were at least 14 h. After reaching equilibrium, GFP fluorescence on individual beads was assessed in a flow cytometer. The RNA dependence and specificity of the binding reactions were assessed as follows to ascertain the validity of the approach.

The observed GFP fluorescence on the beads is RNA-dependent as illustrated in Fig. 2*A*. Background U1A-GFP binding to the beads is low, and the peaks are broad indicative of

a relatively heterogeneous population (Fig. 2, *panel 2*). The observed GFP fluorescence intensity increases with higher U1A RNA concentrations, and the population is more homogenous as indicated by the narrower peak width (Fig. 2, *panels 3* and *4*). The observed signals were reproducible with different RNA and protein preparations.

The assay accurately distinguishes the U1A RNA binding site. Oligonucleotides complementary to the binding site compete with U1A-GFP for the same RNA sequence and thereby reduce observed GFP fluorescence. Only nonspecific background GFP signal is observed for oligos complementary to the binding site (Fig. 2, compare *panels 2* and *8*). On the other hand, oligos complementary to sequences not part of the U1A binding site show significant GFP signal (oligos 2.17 and 3.15, Fig. 2*A*, *panels 3–6*). These oligonucleotides are hybridizing to the RNA without interfering with U1A-GFP binding thereby allowing the observation of the RNA-protein interaction. As a control, oligonucleotides not complementary to the RNA show background nonspecific signal (Fig. 2, compare *panels 2* and *7*).

Decreasing the oligonucleotide length lowers the GFP fluorescence intensity. 20-mers, 17-mers, and 15-mers all yield significant fluorescence, while 10-mers complementary to the same region do not (data not shown). Lower fluorescence is consistently observed for 15-mers, such as oligo 3.15, compared with 17-mers, such as oligo 2.17 (Fig. 2, compare *panels 3* and *4* to *panels 5* and *6*). These observations are not limited to a particular region of the RNA or sequence.

Discrimination of Oligonucleotide Mismatches—The bead assay discriminates between oligonucleotides that contain mismatches under conditions that preserve RNA-protein interactions. When mismatches in the middle of the complementary sequence are introduced, the oligonucleotide yields significantly lower GFP fluorescence. Mismatch discrimination is not unique to a particular sequence as oligonucleotides complementary to distinct regions show a significant difference in GFP fluorescence (Fig. 3*a*, compare oligos 2.17 and 2C.17 and oligos 3.17 and 3C.17). Interestingly, unlike 15-mers and 17-mers, 20-mers do not discriminate mismatches as well (Fig. 3*a*, compare oligos 2.20 and 2C.20 with oligos 2.17 and 2C.17). Also mismatches at the first or second position of either end of the oligonucleotide are not discriminated as well as those in the middle of the sequence (data not shown).

To verify the observed oligonucleotide mismatch discrimination, a compensatory mutation in the RNA was constructed. Binding reactions were prepared with the two different RNAs, and beads coupled to either oligo 2.17 or 2C.17 were added. As observed previously for the wild-type U1A RNA, oligo 2.17 shows significant GFP fluorescence, while oligo 2C.17 does not (Fig. 3*b*). However, the compensatory mutation in the U1A RNA, 77C→G, creates a mismatch for oligo 2.17 and disrupts the hybridization thereby reducing the observed GFP fluorescence. Meanwhile, significant GFP fluorescence is observed for oligo 2C.17, which is complementary to 77C→G RNA (Fig. 3*b*). Compensatory RNA mutations and subsequent U1A binding have also been performed with oligos 3.17 and 3C.17 (data not shown). These "rescue" experiments further verify the observed point mutant hybridization discrimination.

The observed mismatch discrimination is enhanced by measuring hybridization through U1A-GFP binding. In the absence of U1A-GFP, poor discrimination between oligonucleotide mismatches is observed (Fig. 3*c*, compare *panels D* and *E*). This is consistent with reports of poor hybridization behavior of short oligonucleotide sequence (15, 16). However, in the presence of U1A-GFP, the same oligo 2.17 beads show both a higher GFP and Texas Red fluorescence intensity, while two different mismatches of oligo 2.17 show significantly reduced GFP fluorescence (Fig. 3*c*). Thus, in physiological conditions, mismatch discrimination is observed by monitoring hybridization through an RNA-protein interaction.

Detection of Specific RNA-Protein Interactions—To determine whether the observed fluorescence is accurately reflecting the U1A RNA-protein interaction, an A to G point mutation in the U1A loop (96A→G) known to disrupt binding was tested (14). This mutation severely reduces the observed GFP signal as shown in Fig. 4*A*. To quantitate Texas Red and GFP fluorescence, the percentage of the bead population shifted beyond the signals observed for background binding is determined. The quadrants shown in *panels A–C* of Fig. 3*c* determine the cut-offs to define the bead populations with different combinations of Texas Red and GFP fluorescence. Oligos 2.17 and 3.17 show U1A-GFP signal with wild-type U1A RNA, while only background fluorescence is observed with 96A→G. Importantly both oligonucleotides are hybridizing to the RNA as indicated by significant Texas Red fluorescence suggesting that the lower GFP signal is due to disruption of the protein-RNA complex and not reduced hybridization.

As predicted, oligo 1.17, complementary to the U1A binding site, does not show any significant U1A-GFP binding, similar to the experiments described above. Importantly oligo 1.17 is hybridizing to the RNA at levels similar to other oligonucleotides where GFP fluorescence is observed. This suggests that the low observed GFP fluorescence is due to disruption of the RNA-protein complex and not poor hybridization. These data demonstrate the ability of this system to define sequences important for protein recognition on the RNA by footprinting.

To further demonstrate that the observed GFP fluorescence accurately reflects the RNA-protein interaction, the affinity of the U1A complex was measured (Fig. 4*B*). A 75 nM dissociation constant in 300 mM KCl at 25 °C is estimated by curve fitting to a Langmuir isotherm, consistent with published data (14). Meanwhile, the 96A→G point mutant shows no significant binding under the same conditions, consistent with its ~1000-fold weaker affinity for this U1A construct (14). Higher nonspecific U1A-GFP binding to the beads causes broader bead population distributions and is probably responsible for the larger error bars observed at higher protein concentrations.

The observed GFP fluorescence is also sensitive to the salt

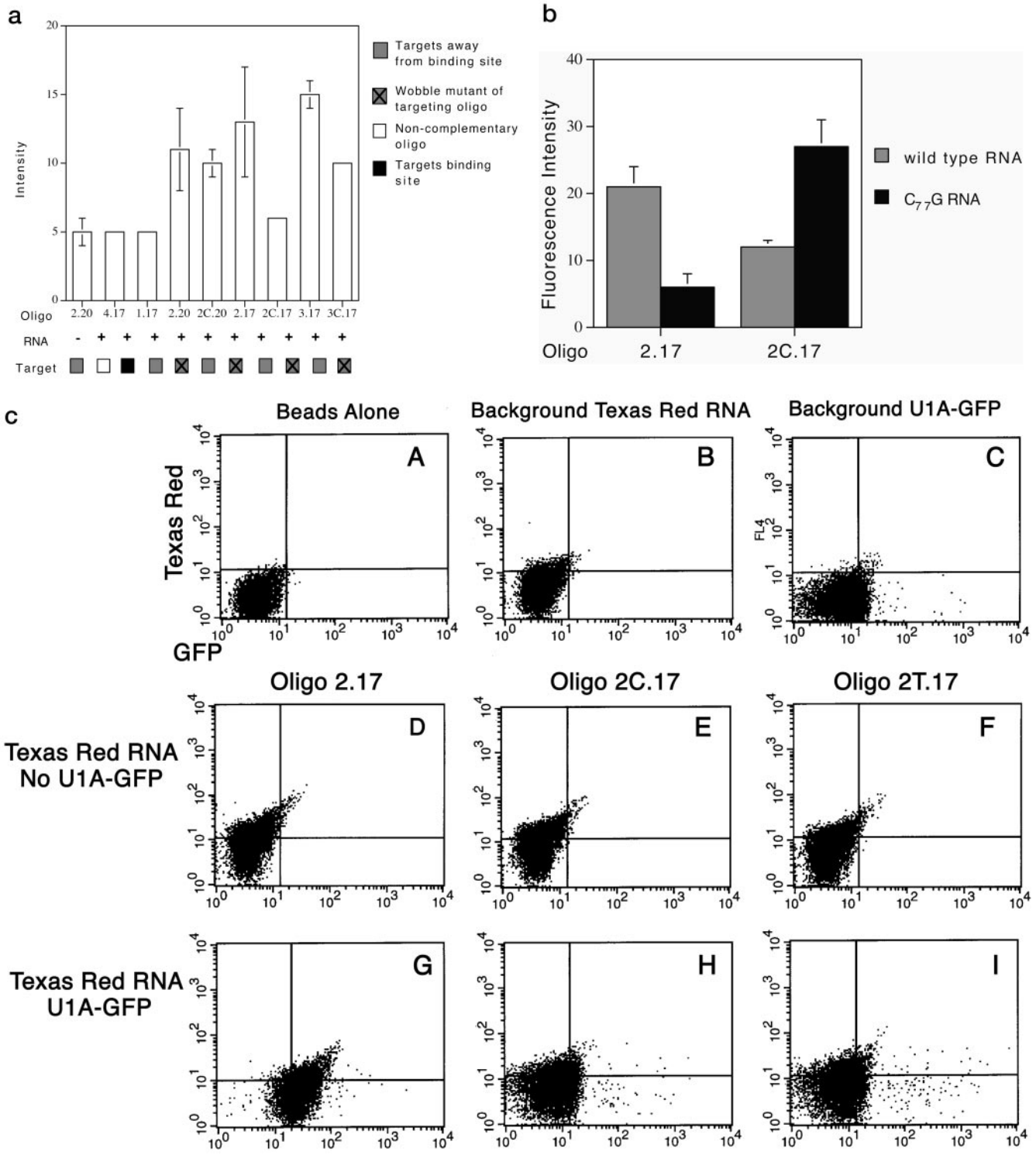

FIG. 3. **Mismatches are discriminated when monitoring hybridization indirectly through U1A-GFP binding.** *a*, the bar graph shows the mean fluorescence of observed GFP signal of bead populations. Fluorescence intensity is in arbitrary units. Triplicate data are averaged, and the *error bars* represent standard deviations. When the standard deviation is less than 1, no error bar is shown. Binding reactions include 75 nM U1A-GFP and 6 nM U1A test RNA. *b*, compensatory RNA mutations restore U1A-GFP binding for an oligonucleotide point mutant. Oligo 2.17 hybridizes to the RNA and shows U1A-GFP binding, while 2C.17 does not. 77C→G, which is an exact match to oligo 2C.17, disrupts hybridization to oligo 2.17 and allows hybridization to oligo 2C.17. Binding reactions include 10 nM U1A RNA and 100 nM U1A-GFP. *Error bars* represent standard deviations of duplicates. *c*, plots show Texas Red fluorescence on the *y* axis and GFP fluorescence on the *x* axis. *Panels A–C* show the nonspecific binding of Texas Red RNA and U1A-GFP to the beads. *Panels D–F* illustrate the poor discrimination of 17-mer oligos

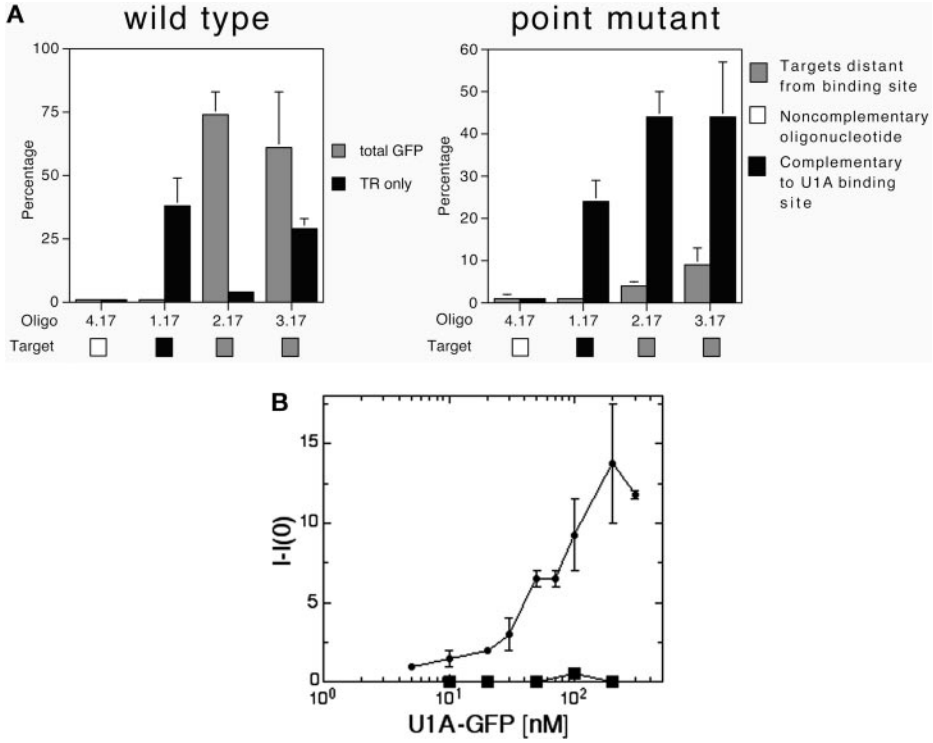

FIG. 4. **Bead assay accurately reflects the RNA-protein interaction.** *A*, percentage of beads shifted above background are tabulated into different categories: beads with Texas Red fluorescence only and beads with GFP fluorescence. U1A-GFP does not bind the 96A→G point mutant RNA, but oligonucleotides still significantly hybridize to the RNA as indicated by the Texas Red fluorescence. *Error bars* represent standard deviations of triplicate experiments. *B*, titration of U1A-GFP into 3 nM RNA generates binding curves. *Squares* represent the 96A→G point mutant, and *circles* represent the wild-type U1A RNA. Background fluorescence (I_o) in the absence of RNA is subtracted for each protein concentration. Binding curves were performed in duplicate and averaged. *Error bars* represent standard deviations. Curve fitting suggests U1A-GFP is binding with a dissociation constant of ~75 nM in 300 mM KCl at 25 °C.

concentration. At higher KCl concentrations, the GFP fluorescence intensity increases by ~25% at each protein concentration. However, the dissociation constant shifts from ~35 to ~75 nM. These observations reflect stronger hybridization and weaker U1A interaction at higher salt concentrations. In sum, these data demonstrate the specificity and affinity of the U1A RNA-protein interaction on beads.

Specific Binding in Mixed Populations—For screening RNA-protein interactions with the bead assay, RNAs will have to be identified from complex mixtures of RNAs. To determine whether total RNA can compete with U1A-GFP binding, yeast RNA was added to the binding reactions. Human U1A does not specifically bind any yeast RNA (17, 18). Only in the presence of the U1A test RNA is GFP fluorescence detectable with oligo 2.17 as shown in Fig. 5*A*. Oligonucleotides not complementary to the U1A RNA such as oligo 4.17 show fluorescence equivalent to background. This suggests that even in contexts where there might be high nonspecific binding, specific binding is observable. These experiments were performed with larger 5.7-μm diameter beads because the sensitivity is higher (data not shown). The nonspecific binding on these larger beads is also higher as more GFP fluorescence is observed. However, specific GFP fluorescence is observed at lower concentrations of U1A-GFP compared with the smaller 2.8-μm microbeads.

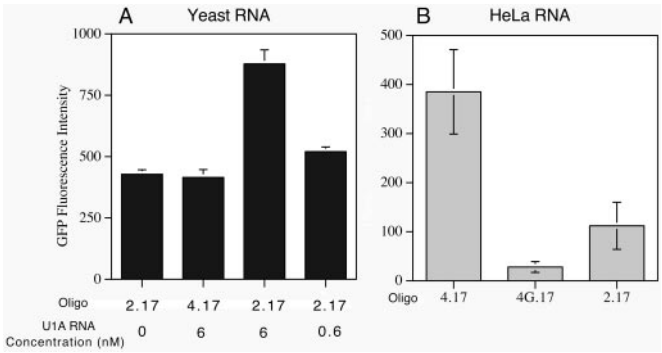

FIG. 5. **U1A-GFP binding to the U1A test RNA and to U1 snRNA is detectable in total RNA backgrounds.** *A*, oligo 2.17 but not oligo 4.17 detects U1A-GFP signal in 0.1 μg/μl yeast total RNA and 25 nM U1A-GFP. Fluorescence intensities are higher with larger 5.7-μm beads. Data are in triplicate with standard deviation *error bars*. *B*, conversely oligo 4.17 shows GFP fluorescence in HeLa total RNA. U1A-GFP is binding to snRNA in the total HeLa RNA. Oligo 4.17 also shows higher GFP fluorescence than a mismatch, oligo 4G.17, or a noncomplementary sequence, oligo 2.17. Reactions include 0.16 μg/μl HeLa total RNA (40 μg total) and 50 nM U1A-GFP. *Error bars* represent standard deviations of triplicates. Background signals were subtracted for this graph.

To determine whether RNAs isolated from RNA preparations can be identified with microbeads, total RNA was isolated from HeLa cells and mixed with U1A-GFP. In HeLa RNA,

binding to the RNA in the absence of U1A-GFP. Oligonucleotide mismatches do not significantly affect the observed Texas Red signal. Only the wild-type oligonucleotide shown in *panel G* gives significant GFP fluorescence with 100 nM U1A-GFP and 30 nM U1A RNA, while the mismatches, in *panels H* and *I*, show some Texas Red signal but no significant GFP fluorescence.

U1A binds to U1 snRNA and its own mRNA. An oligonucleotide complementary to snRNA, oligo 4.17, shows higher GFP fluorescence compared with a mismatch, oligo 4G.17, and a noncomplementary oligonucleotide, oligo 2.17, as illustrated in Fig. 5*B*. The U1 snRNA concentration is ~1–5 nM in these experiments. The observed GFP fluorescence is HeLa RNA-dependent as the signal varies with HeLa RNA concentration but remains unchanged with increasing equivalent concentrations of yeast RNA. In sum, these data demonstrate that a specific RNA can be identified with microbeads from total RNA.

An additional requirement for screening with oligonucleotide bead libraries is the ability to detect a small percentage of oligonucleotide beads from a large background of beads that do not bind. Initial experiments diluting the oligonucleotide beads 100-fold suggest that at even low oligonucleotide concentrations the RNA-protein complex can be detected. Further experiments diluting oligo 2.17 in a large excess of oligo 4.17 demonstrated that two populations of beads could be differentiated by GFP fluorescence (data not shown). These data also demonstrate that the observed shifts are sufficient to identify an RNA-protein complex.

DISCUSSION

We describe a system to monitor RNA-protein interactions in solution with microbeads using flow cytometry. We demonstrate the versatility of the approach for 1) discriminating between mismatches in the oligonucleotides, 2) mapping protein recognition sites on RNA, 3) differentiating specific and nonspecific binding RNAs, and 4) detecting specific RNAs in complex mixtures. Importantly, specific binding can be detected in high nonspecific RNA backgrounds, and the system can discriminate a nonspecific binding point mutant at a variety of protein and RNA concentrations. Because flow cytometry is used to monitor the fluorescence on distinct oligonucleotide-coupled microbeads, the system is amenable to high throughput, genomic scale screening of RNA-protein interactions.

With the U1A interaction, we have determined the fundamental requirements for using this microbead system for screening RNA-protein interactions. The approximate 2-fold changes in observed GFP fluorescence are sufficient to distinguish specific protein binding to RNA from background. Future versions of the system may have increased sensitivity and dynamic range by using brighter fluorophores and microbeads with higher oligonucleotide densities. Furthermore, because the system is monitoring the protein binding to a distribution of thousands of microbeads, the fluorescence shifts are more significant than just monitoring the bulk signal. Methods to analyze the distributions more quantitatively are being developed. The microbead system also allows for the measurement of relative affinities of a protein for its cognate RNA thereby distinguishing specific and nonspecific binding during screening and thereby reducing false positives. Two approaches are possible. After initial screening at a particular protein concentration, binding experiments will help distinguish specific and nonspecific binding candidates. Alternatively, screening at different protein concentrations could be performed to determine relatively strong and weak binding interactions.

The microbead-based system described here has a number of advantages over other recently developed RNA-protein screening strategies. It is rapid with few time-consuming manipulations required. Also, unlike many *in vivo* strategies, there are no limitations to the size of the RNA or its basic structural features. It is often difficult to monitor the binding of large RNAs directly because they do not migrate well in gel electrophoresis for mobility shift studies. With the microbead system, the only requirement is that oligonucleotides hybridize to regions distant from the binding site. Most mRNAs have multiple regions accessible to hybridization in physiological conditions (19).

With the system presented here, hybridization is monitored indirectly through the protein fluorescence. To observe protein fluorescence, the oligonucleotides need to hybridize to particular RNAs that are also binding the protein. The combination of observed RNA hybridization and protein fluorescence on a microbead indicates that an RNA-protein complex is present. Thus, many RNAs may be hybridizing to the beads, but only when the fluorescently labeled protein is bound with high affinity to one of these RNAs is positive signal observed. This significantly reduces the nonspecific binding that would be observed in identifying possible RNA targets compared with other strategies that isolate all the RNAs bound to beads. In its current form, the system does not reach equilibrium for hours presumably because of slow hybridization at physiological temperatures. Smaller volumes may help reduce the time required to reach equilibrium.

The mismatch discrimination observed with this bead strategy may allow it to be adopted for single nucleotide polymorphism analysis. Similar to assays such as the invasive cleavage method (20), monitoring hybridization indirectly provides a sensitivity enhancement to observe the subtle effects of mismatches on hybridization. The enhanced mismatch discrimination observed through the protein interaction may be particular to the U1A system. Future studies of other RNA-binding proteins will determine the generality of the observed mismatch discrimination.

For genomic screening, proteins bound to RNA could be challenged with large oligonucleotide coded bead libraries. The coded beads would be sorted in a flow cytometer while monitoring RNA and protein fluorescence to determine which sequences are hybridizing to the RNA while preserving the RNA-protein complex. This information can then be compared with sequenced genomes to determine which RNAs are binding and which sequences may be important for the interaction. Various coding strategies are currently being developed that do not require decoding or very large beads (9–12).

In sum, we have developed a microbead-based system to determine which RNAs may be binding a particular protein as well as which RNA sequences may be important for the RNA-protein interaction. Many applications of the assay are possible including binding in cell extracts, single nucleotide polymorphism analysis, and monitoring the effects of small molecules on RNA-protein complexes. Perhaps the most inviting aspect of this system is to use large coded oligonucleotide bead libraries to probe RNA-protein interactions on genomic scales.

Acknowledgments—We thank Kathy Hall for U1A plasmids, Vicki Losick for constructing 96A→G, and the staff at the Dana-Farber Cancer Institute flow cytometry facility. We thank M. Damelin, C. Stern, P. Zarrinkar, and D. Treiber for critical reading of the manuscript and Jeff Way and Matt Trau for stimulating discussions.

* A portion of this work was supported by grants from the National Institutes of Health and the Department of Defense (to P. A. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by the Claudia Barr Young Investigator program.

§ To whom correspondence should be addressed: Dana-Farber Cancer Inst., SM 922, 1 Jimmy Fund Way, Boston, MA 02115. Tel.: 617-632-5105; Fax: 617-632-5103; E-mail: alex_brodsky@dfci.harvard.edu.

REFERENCES

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921
2. Zhang, B., Kraemer, B., SenGupta, D., Fields, S. & Wickens, M. (2000) Yeast three-hybrid system to detect and analyze RNA-protein interactions. *Methods Enzymol.* **318,** 399–419
3. Danner, S. & Belasco, J. G. (2001) T7 phage display: a novel genetic selection system for cloning RNA-binding proteins from cDNA libraries. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 12954–12959
4. Paraskeva, E. & Hentze, M. W. (2000) Translational repression assay procedure: a method to study RNA-protein interactions in yeast. *Methods Enzymol.* **318,** 374–384
5. Andrews, L. G. & Keene, J. D. (1999) Identification of specific protein-RNA target sites using libraries of natural sequences. *Methods Mol. Biol.* **118,** 233–244
6. Tenenbaum, S. A., Carson, C. C., Lager, P. J. & Keene, J. D. (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 14085–14090
7. Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., Jin, X., Feng, Y., Wilkinson, K. D., Keene, J. D., Darnell, R. B. & Warren, S. T. (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107,** 477–487
8. Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T. & Darnell, R. B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* **107,** 489–499
9. Battersby, B. J., Lawrie, G. A. & Trau, M. (2001) Optical encoding of microbeads for gene screening: alternatives to microarrays. *Drug Discov. Today* (*HTS Suppl.*) **6,** S19
10. Chan, W. C., Maxwell, D. J., Gao, X., Bailey, R. E., Han, M. & Nie, S. (2002) Luminescent quantum dots for multiplexed biological detection and imaging. *Curr. Opin. Biotechnol.* **13,** 40–46
11. Cao, Y. C., Jin, R. & Mirkin, C. A. (2002) Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. *Science* **297,** 1536–1540
12. Cunin, F. *et al.* (2002) Biomolecular screening with encoded porous-silicon photonic crystals. *Nat. Mater.* **1,** 39–41
13. Wilkins Stevens, P., Henry, M. R. & Kelso, D. M. (1999) DNA hybridization on microparticles: determining capture-probe density and equilibrium dissociation constants. *Nucleic Acids Res.* **27,** 1719–1727
14. Zeng, Q. & Hall, K. B. (1997) Contribution of the C-terminal tail of U1A RBD1 to RNA recognition and protein stability. *RNA* **3,** 303–314
15. Guo, Z., Guilfoyle, R. A., Thiel, A. J., Wang, R. & Smith, L. M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* **22,** 5456–5465
16. Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. & Brown, E. L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14,** 1675–1680
17. Tang, J. & Rosbash, M. (1996) Characterization of yeast U1 snRNP A protein: identification of the N-terminal RNA binding domain (RBD) binding site and evidence that the C-terminal RBD functions in splicing. *RNA* **2,** 1058–1070
18. Brodsky, A. S. & Silver, P. A. (2000) Pre-mRNA processing factors are required for nuclear export. *RNA* **6,** 1737–1749
19. Ho, S. P., Bao, Y., Lesher, T., Malhotra, R., Ma, L. Y., Fluharty, S. J. & Sakai, R. R. (1998) Mapping of RNA accessible sites for antisense experiments with oligonucleotide libraries. *Nat. Biotechnol.* **16,** 59–63
20. Wilkins Stevens, P., Hall, J. G., Lyamichev, V., Neri, B. P., Lu, M., Wang, L., Smith, L. M. & Kelso, D. M. (2001) Analysis of single nucleotide polymorphisms with solid phase invasive cleavage reactions. *Nucleic Acids Res.* **29,** E77
21. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288,** 911–940