Biol. Chem., Vol. 382, pp. 1157 – 1162, August 2001

Binding of IRE-BP to Its Cognate RNA Sequence: SFM Studies on a Universal RNA Backbone for the Analysis of RNA-Protein Interaction

Michael Bonin[1], Jürgen Oberstrass[1], Ulrike Vogt[2], Michael Wassenegger[2] and Wolfgang Nellen[1,*]

[1] Abteilung Genetik, Universität Kassel, Heinrich-Plett-Str. 40, D-34132 Kassel, Germany
[2] Fraunhofer-Institut für Umweltchemie und Ökotoxikologie, Abt. Molekulare Biotechnologie, Am Klopferspitz 18A, D-82152 Martinsried, Germany

* Corresponding author

We have used an RNA consisting of the potato spindle tuber viroid (PSTVd) and 240 bp of double-stranded RNA derived from the GUS gene as a backbone for scanning force microscope (SFM) studies on RNA binding proteins. The *in vitro* transcribed RNA forms a rod-like structure of apparent 130 nm in length with a completely base paired central part flanked by the incompletely paired viroid helix with bulges on both sides. The termini of the molecule consist of loops such that no blunt or staggered RNA ends are exposed. Suitable, asymmetrical restriction sites in the construct allow for the insertion of sequences of interest, *e. g.* protein binding sites. We have inserted the IRE (iron responsive element) sequence into the construct and have used *in vitro* transcripts to study binding of IRE-BP. Relative binding frequencies show that 70% of the protein binds to the expected site in the molecule while only a slightly enhanced binding is observed at the termini. In the GUS-PSTVd-IRE backbone, the orientation of the molecule is easily determined by IRE-BP binding. It thus provides a versatile tool to study specific as well as preferential interaction of other proteins with sequences or structures inserted into a different part of the molecule.
Key words: AFM / dsRNA / Single molecule analysis.

Introduction

The analysis of protein–nucleic acid interaction is mostly done by biochemical methods. The specificity of binding sites, affinity and potential structural changes in the target nucleic acid can thus be determined. In recent years, however, single molecule analysis has allowed for visualization of RNA-protein complexes by SFM (Lyubchenko *et al.*, 1995; Smith *et al.*, 1997; Keller *et al.*, 2000), the measurement of binding forces (De Paris *et al.*, 2000) and the identification of preferential binding sites of 'promiscuous' binding proteins (Bonin *et al.*, 2000). The observation that proteins may have a general affinity to a DNA or RNA substrate but display a preference for specified sequence motifs can so far not be easily approached by biochemistry, especially when the motif is not defined. In contrast, SFM imaging techniques give sufficient resolution on an extended nucleic acid molecule to identify binding regions. The evaluation of many molecules gives a probability of binding along the length of the target and thus identifies preferred binding regions. Molecular biology techniques, such as site-directed mutagenesis within a 40 bp region, may then define specific sequence requirements. An inherent problem of the identification of binding regions is that the orientation of the nucleic acid is not known. Although this can be solved by the cumbersome analysis of subfragments (Bonin *et al.*, 2000), we intended to design an oriented backbone specifically for RNA binding proteins which could be generally applicable and allow for a comparison of specific and preferred binding to a defined site with promiscuous binding to a complete dsRNA helix and an incomplete bulged helix. A double-stranded or partly double-stranded RNA (dsRNA) molecule is advantageous because it is stiff and rod-like, while single-stranded RNA forms large tangles in SFM images. In addition, dsRNA is resistant to many nucleases and therefore less sensitive to the manipulations required for SFM. Thus, the PSTVd-GUS hybrid molecule could serve very well for these investigations. The two modules of the molecule consist of the potato spindle tuber viroid (PSTVd) RNA and part of the β-glucuronidase gene from *Escherichia coli*. PSTVd is a plant pathogen which consists of only a 359 nt long single-stranded RNA.

IRE-BP is one of the best defined RNA binding proteins. It functions in iron homeostasis by regulating the half life of the ferritin mRNA and the translation of the transferrin receptor mRNA (Paraskeva and Hentze, 1996). The 98 kD protein binds specifically to a bulged stem-loop with a K_D of 0.25 nM (Zheng *et al.*, 1992). Our experiments addressed the following questions:

(i) is specific IRE-BP binding to its cognate sequence detected under SFM conditions?
(ii) Is binding at low protein concentrations sufficiently frequent to obtain a large number of complexes and thus allow for the determination of the orientation of the RNA molecule?
(iii) Is the previously observed end-binding of proteins

(Goodman *et al.*, 1983; Michalowski *et al.*, 1999; Bonin *et al.*, 2000) abolished or reduced by using 'closed termini' (*i. e*. terminal loops instead of open 3' and 5' ends)?

If these conditions were met, the GUS-PSTVd-IRE backbone could be used as a general tool to investigate RNA-protein binding in an oriented molecule.

Results

The secondary structure of the PSTVd-GUS (PG) hybrid molecule with the inserted IRE stem-loop (PG-I) was calculated using the Zuker program (Zuker *et al.*, 1999; Mathews *et al.*, 1999). As shown in Figure 1, the right PSTVd arm corresponds to the viroid structure (Riesner *et al.*, 1979), the structure prediction of the left arm is slightly different from the standard PSTVd secondary structure due to the introduction of two restriction sites and the junction with the inverted-repeat GUS sequence. The IRE stem-loop can be seen on the right. The IRE-BP structure as proposed by Brookhaven database 1NIS (Lauble *et al.*, 1994) is modeled on the stem loop. According to this prediction, the RNA molecule has an overall length of 130 nm (approx. 84% of the molecules were between 120 and 130 nm, approx. 97% were between 115 and 130 nm) and IRE-BP has a diameter of 11 – 13 nm.

Binding assays of IRE-BP and *in vitro* transcribed PG-I RNA were performed with various protein to RNA ratios deduced from gel-shift experiments (data not shown). The optimal ratio as determined by SFM inspection was 10 proteins/1 RNA molecule (see Material and Methods) and yielded a background count of 20 free PG-I RNA molecules and 40 free IRE-BP molecules per 100 molecules of the complex. RNA molecules with two bound proteins were extremely rare and constituted about 0.5% of the complexes. An overview and representative single molecules are shown in Figure 2. As previously observed by others (Schaper and Jovin,1996) and by us (Bonin *et al.*, 2000), the dimensions of the RNA molecule differ from the expected values: the height was 1.5 nm and the width was 5 – 7 nm instead of 2.5 nm in both dimensions. The IRE-BP had a diameter of 7 – 10 nm and a height of 2.5 nm.

RNA molecules in the images were subdivided into 10 segments of 13 nm and bound protein was assigned to these segments. Since the orientation of the RNA could not be determined by its structure in the SFM image, 'left and top' were initially arbitrarily assigned to the left end of the RNA while 'right and bottom' were assigned to the right end (see Bonin *et al.*, 2000). A symmetrical distribution of approximately 35% of the protein bound to segments 3 and 8 was found (27 to 39 nm from either the left or the right end). The residual 30% of the bound molecules was rather equally distributed with approximately 3% per segment but a slight preference for the ends (Figure 3 and data not shown). We conclude that 70% of the protein was in fact bound to the specific site since no internal binding preference was found in a control molecule (GUS-PSTVd without IRE) and since simultaneous binding of two protein molecules to segments 3 and 8 was never detected. Based on this, the relative binding frequencies shown in Figure 3 were calculated. Although localization within a segment was difficult, we found that IRE-BP was mostly bound to the outer part of the sement, *i. e*. closer to the terminus of the backbone.

With 7%, end binding preference was significantly lower than in previous experiments with other proteins (Bonin *et al.*, 2000) but still significant. To see if end bind-

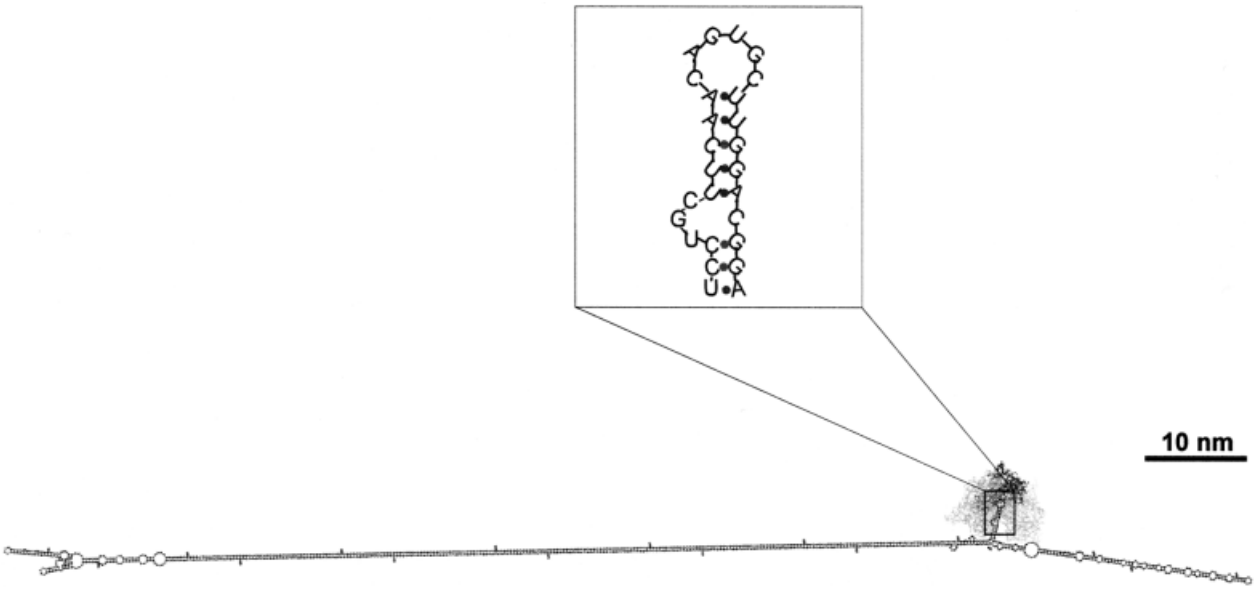

Fig. 1 Calculated Structure of the PG-I Hybrid Molecule.
IRE-BP is modelled onto the cognate sequence. The bar represents 10 nm.

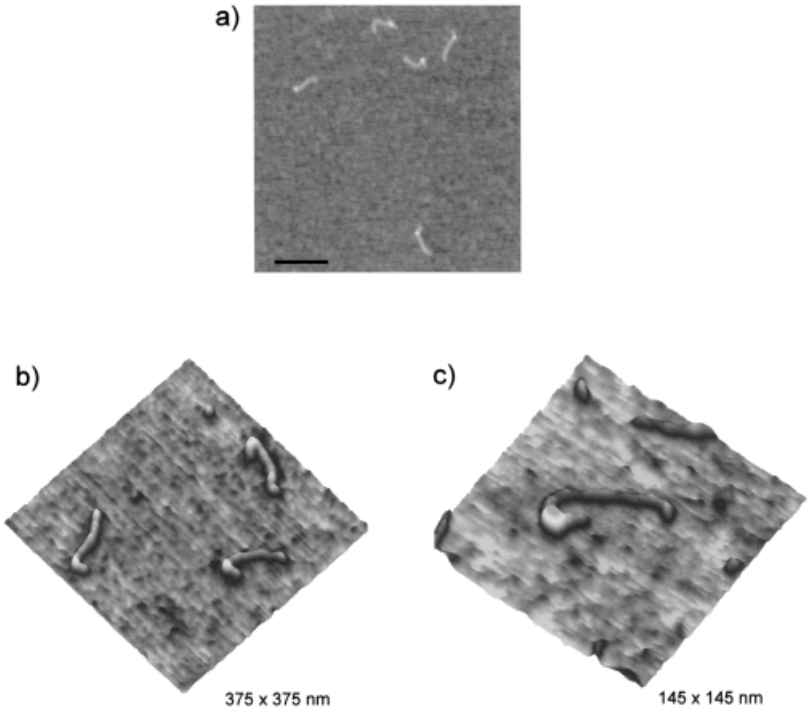

Fig. 2 Overview of a Field of PG-I and IRE-BP Molecules (a) and Surface Plots of PG-I-IRE-BP Complexes (b,c).
(a) Nucleic acid-protein complexes are clearly seen as dots on the linear RNA structure. The molecule on the left shows end binding of IRE-BP while the others show binding to the specific site. The bar represents 200 nm.
(b) Surface plot of different PG-I-IRE-BP complexes. Image size is 375 × 375 nm.
(c) Surface plot of a single PG-I-IRE-BP complex; the average height of the protein is 2 nm. Image size is 145 × 145 nm.

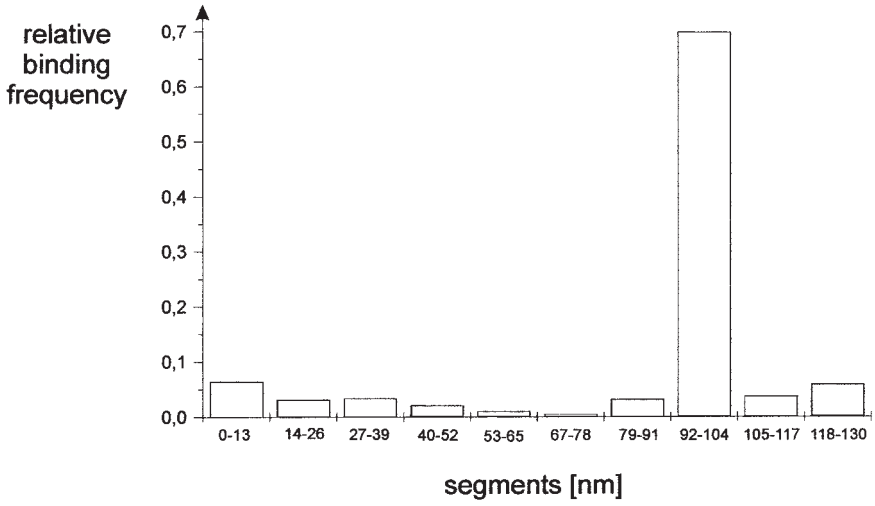

Fig. 3 Relative Binding Frequency of IRE-BP to PG-I as Calculated from 323 Nucleic Acid-Protein Complexes.
The RNA molecule was divided into 10 segments of 13 nm (corresponding to 43bp each) and each bound protein was assigned to a segment. See text for details.

ing was reduced because of the closed termini in the PB-I target we examined non-specific binding of IRE-BP to a substrate with open termini. A 707 bp dsRNA described previously (Hildebrandt and Nellen, 1992) was used for this experiment. End binding preference was calculated by the following procedure: the percentage of protein binding to an end was divided by the average percentage of protein binding to an internal segment. The value was denominated 'end binding preference' (ebp). The factor was calculated from 702 J2 and from 795 K1 anti dsRNA antibodies binding to dsRNA with open termini. Ebp was also determined from 201 and 323 IRE-BP molecules binding to dsRNA with open and closed termini, respectively. For calculating binding to dsRNA with closed termini, the values from Figure 3 were used but the segment containing the specific binding site was eliminated. For

Table 1 End Binding Preference of Proteins to RNA.

RNA substrate	dsRNA mAb J2[a]	dsRNA mAb K1[b]	IRE-BP[c]
Open terminus[d]	5.0	6.25	5.5
Closed terminus[e]	n.d.	n.d.	1.6

[a] Bonin *et al.*, 2000.
[b] Bonin, Klaue, Lukacs and Nellen, unpublished.
[c] This study.
[d] 707 bp dsRNA (Bonin *et al.*, 2000).
[e] PG-I.

End binding preference (ebp) was calculated by the following way: % average binding to end segments (%) was divided by % average binding to internal segments (%) yielding the factor for preferential end binding. For proteins (J2 and IRE-BP) which displayed a significant binding preference to an internal fragment of the RNA, this segment was not included in the calculation.

$$\text{ebp} = \frac{\Sigma \text{ bdg. end seg.}}{(\Sigma \text{ all bdg.}) \times 2 \times 100} : \frac{\Sigma \text{ bdg. int. seg.}}{(\Sigma \text{ all bdg.}) \times \#\text{ int. seg.} \times 100}$$

J2, data from Bonin *et al.* (2000) were used, eliminating the preferential binding sites. The K1 antibody did not show any preferential binding sites except for the ends (Klaue, Bonin and Nellen, unpublished). The data are summarized in Table 1.

Discussion

We have previously shown that SFM is a versatile tool for examining RNA-protein interactions on a single molecule basis and for determining binding preferences which are difficult or impossible to identify by conventional biochemical methods. Though binding sites can be localized accurately within approximately 40 bp, SFM measurements pose the problem that the orientation of the molecule cannot be unambiguously determined. Another drawback is the apparently non-specific end binding preference of proteins with high affinity to nucleic acids. This can obscure the detection of specific binding and may, in extreme cases, titrate out a binding protein when non-sequence-specific end binding preference is higher than the specific interaction. This is especially the case when proteins with sequence preference but not sequence specificity are examined (see *e. g*. Bonin *et al.*, 2000). We have significantly decreased end binding preference by using RNA molecules with 'closed termini', *i. e*. single-stranded RNA which folds back and does not expose the 5' and 3' ends at the termini of the rod structure. Even though the 5' and 3' ends of the molecule (contained in segment 1) are not completely buried in a double helical structure, they are not 'open' since they are not available for hybridization with an oligonucleotide (data not shown). As seen in Table 1, closed termini reduced the ebp from 5–6 to approximately 1.5. The closed termini could possibly have abolished end binding completely and the residual binding observed in our experiments may be a result of some nuclease cleavage in the single-stranded loops at the termini of the RNA molecule.

Binding of IRE-BP was found in the expected segment of the GUS-PSTVd backbone with a slight shift to the distal part of the segment. We assume that this was due to bending of the IRE-loop to the right which, could be caused by the kink in the backbone (see Figure 2).

The GUS-PSTVd backbone provides an excellent general substrate for SFM measurements for several reasons:

(1) the molecule is of sufficient size to unambiguously distinguish at least 10 segments;
(2) due to the high structural stability, the insertion of sequence motifs can be done without major disturbance of the overall structure of the molecule;
(3) in contrast to *in vitro* hybridization of two complementary RNAs, substrates are generated from a single RNA molecule with high efficiency and sequence motifs of interest are exposed from a stiff and stable rod-like structure;
(4) since the complementary parts of the GUS gene are separated by PSTVd sequences, recombination is not a significant problem in cloning;
(5) the completely double-stranded part of GUS and the partially double-stranded part of PSTVd provide a good internal control for binding specificity to a sequence of interest.

This last point is especially of interest since many RNA binding proteins have a general affinity for RNA with an individual preference for dsRNA, ssRNA or incomplete hybrids. Since all these features are contained in the backbone, this allows for a direct and quantitative comparison of sequence specific and sequence-non-specific binding. Similar to biochemical assays, the affinity of a protein to mutated sites can be determined. Quantitative examination will, however, be more reliable and exact when both wild type and mutant are inserted into the same backbone. As can be seen in Figure 2, structural features in the nucleic acid (like bending and kinking) can be readily seen by SFM imaging while it takes laborious cloning efforts to investigate this by biochemical meth-

ods. In this case, however, the kink was most likely not induced by protein binding since it was also observed in RNA without bound protein.

The insertion of the IRE motif and the highly specific and efficient binding we have described here, allow for determination of the orientation of the molecule. A binding site for a different protein inserted into the left arm of the backbone can be readily examined relative to IRE-BP binding, provided the second binding protein can be distinguished by size from IRE-BP.

We have not yet addressed the question whether binding can be observed under SFM conditions; the GUS-PSTVd backbone will, however, most likely provide an excellent substrate to investigate RNA-Protein binding *in situ* in a fluid-cell.

Materials and Methods

DNA Templates

Complementary single-stranded oligodeoxyribonucleotides were synthesized by Life Technologies (Karlsruhe, Germany). The synthetic IRE contains *Bgl*II linker sites with the sequence from 5´ to 3´: GATCCTGCTTCAAGAGTGCTTGGACG (for ferritin IRE). The complementary oligos were annealed and filled in with Klenow fragment I and 5´phosphorylated with T4 polynucleotid kinase, according to standard procedures (Maniatis *et al.*, 1982).The plasmid pT3T7gus-PSTVd (Vogt *et al.*, unpublished) contains a single *Sma*I restriction site. The annealed modified oligonucleotides were cloned into this *Sma*I site using T4 ligase and the restriction enzyme controlled ligation. Recombinant plasmids were transformed into DH5α *E. coli* cells. Plasmids were selected by restriction enzyme cleavage and confirmed by sequencing, the resulting construct was designated pT3T7gus-PSTVd-IRE.

RNAs and Proteins

The GUS-PSTVd-IRE RNA (PG-I) was synthesized by *in vitro* transcription with T7 RNA polymerase from plasmids linearized with *Sal*I. The RNA was purified on an 8% polyacrylamid/8 M urea gel, eluted, and precipitated with ethanol.

Recombinant human IRP1 was purified from the *E. coli* strain BL21 DE3-pT7-his-hIRF by affinity chromatography on Ni^{2+}-NTA beads (Qiagen, Hilden, Germany) as described previously (Gray *et al.*, 1993). The protein was eluted with 50 mM imidazol, dialyzed against 50 mM Hepes-NaOH, pH 7.6, 150 mM KOAc, 1.5 mM $MgCl_2$, 5% glycerol and stored in small aliquots at –80 °C.

J2 and K1 are monoclonal antibodies directed against dsRNA (Schönborn *et al.*, 1991; Lukacs, 1994), while J2 has a preferential binding site to the sequence motif A2N9A3N9A2, no preferential binding site was detected with K1. Both antibodies have strong, sequence-unspecific binding preference to open termini (Bonin *et al.*, 2000; Bonin, Klaue, Lukacs and Nellen, unpublished).

Sample Preparation and Scanning Force Microscopy

Gus-PSTVd-IRE-IRE-BP complexes were formed by combining 1.0 – 2.5 nM dsRNA and 20 nM IRE-BP in 10 mM Hepes, pH 7.6, 3 mM $MgCl_2$, 40 mM KCl, for 30 minutes at 0 °C. A 10 µl drop of the sample was placed for 1 min on a freshly cleaved, glow discharged (1 min, 0.3 mbar air) mica surface, washed with 1 ml MilliQ water and blown dry with nitrogen. Samples were scanned with a Nanoscope III multimode SFM (Digital Instruments, Santa Barbara, USA) operated in the tapping mode using a J scanner with a 125 × 125 (x,y) × 5 (z) µm scan range or an E scanner with a 10 × 10 (x,y) × 2.5 (z) µm scan range. Microfabricated silicon tips with a force constant of 42 N m^{-1} and a resonance frequency of 320 kHz (Pointprobes, Nanosensors GmbH, Wetzlar, Germany) were used. Images were obtained in the topographic mode. Measurements were done in air (relative humidity 20 – 50%) at room temperature (18 – 27 °C). Images (512 × 512 pixels) were taken at 1 – 2 Hz scanning frequency.

Image Analysis

Images were processed with the Nanoscope software including the operations of plane fitting, flattening and geometric filtering. Contour length measurement was done by drawing a traverse line along the skeleton of the molecule on the screen.

For quantitative analysis of IRE-BP binding to the RNA molecules, molecules were subdivided into segments of 13 nm. Left and down was arbitrarily defined as segment one. The number of IRE-BP found in each segment was counted and divided by the total number of bound proteins (relative binding frequency).

Acknowledgements

We thank Dr. M. Hentze for providing the expression vector for IRE-BP. Dr. R. Kassing is acknowledged for providing access to the DI Nanoscope. M.B. was supported by a stipend from the Otto-Braun-Fonds.

References

Bonin, M., Oberstrass, J., Lukacs, N., Ewert, K., Oesterschulze, E., Kassing, R., and Nellen, W. (2000). Determination of preferential binding sites for anti-dsRNA antibodies on double-stranded RNA by scanning force microscopy. RNA *6*, 563 – 570.

De Paris, R., Strunz, T., Oroszlan, K., Güntherodt, H.-J., and Hegner, M. (2000). Force spectroscopy and dynamics of the biotin avidin bond studied by scanning force microscopy. Single Molecules *1*, 285 – 290.

Goodman, T.C., Nagel, L., Rappold, W., Klotz, G., and Riesner, D. (1984). Viroid replication: equilibrium association constant and comparative activity measurements for the viroid-polymerase interaction. Nucleic Acids Res. *12*, 6231 – 6246.

Gray, N.K., Quick, S., Goossen, B., Constable, A., Hirling, H., Kuhn, L.C., and Hentze, M.W. (1993). Recombinant iron-regulatory factor functions as an iron-responsive-element-binding protein, a translational repressor and an aconitase. A functional assay for translational repression and direct demonstration of the iron switch. Eur. J. Biochem. *218*, 657 – 67.

Hildebrandt, M., and Nellen, W. (1992). Differential antisense transcription from the *Dictyostelium* EB4 gene locus: Implications on antisense mediated regulation of mRNA stability. Cell *69*, 197 – 204.

Keller, R.W., Kuhn, U., Aragon, M., Bornikova, L., Wahle, E., and Bear, D.G. (2000). The nuclear poly(A) binding protein, PABP2, forms an oligomeric particle covering the length of the poly(A) tail. J. Mol. Biol. *297*, 569 – 583

Lauble, H., Kennedy, M.C., Beinert, H., and Stout, C.D. (1994). Crystal structures of aconitase with *trans*-aconitate and nitrocitrate bound. J. Mol. Biol. *237*, 437 – 451.

Lukacs, N. (1994). Detection of virus infection in plants and dif-

ferentiation between coexisting viruses by monoclonal antibodies to double-stranded RNA. J. Virol. Methods *47,* 255–272.

Lyubchenko, Y.L., Jacobs, B.L., Lindsay, S.M., and Stasiak, A. (1995). Atomic force microscopy of nucleoprotein complexes. Scanning Microsc. *9*, 705–727.

Maniatis, T, Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning. A Laboratory Manual (Cold Spring Harbor, New York, USA: Cold Spring Harbor Laboratory Press).

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded Sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. *288*, 911–940.

Michalowski, S., Miller, J.W., Urbinati, C.R., Paliouras, M., Swanson, M.S., and Griffith, J. (1999). Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. Nucleic Acids Res. 27, 3534–3542.

Paraskeva, E., and Hentze, M.W. (1996). Iron-sulphur clusters as genetic regulatory switches: the bifunctional iron regulatory protein-1. FEBS Lett. *389*, 40–43.

Riesner, D., Henco, K., Rokohl, U., Klotz, G., Kleinschmidt, A.K., Gross, H.J., Domdey, H., Jank, P., and Sänger, H.L. (1979). Structure and structure formation of viroids. J. Mol. Biol. *133*, 85–115.

Schaper, A., and Jovin, T.M. (1996). Striving for atomic resolution in biomolecular topography: the scanning force microscope (SFM). Bioessays *18*, 925–935.

Schönborn, J., Oberstrass, J., Breyel, E., Tittgen, J., Schumacher, J., and Lukacs, N. (1991). Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. Nucleic Acids Res. *19*, 2993–3000.

Smith, B.L., Gallie, D.R., Le, H., and Hansma, P.K. (1997). Visualization of poly(A)-binding protein complex formation with poly(A) RNA using atomicforce microscopy. J. Struct. Biol. *119*, 109–117.

Zheng, L., Kennedy, M.C., Blondin, G.A., Beinert, H., and Zalkin, H. (1992). Binding of cytosolic aconitase to the iron responsive element of porcine mitochondrial aconitase mRNA. Arch. Biochem. Biophys. *299*, 356–360.

Zuker, M., Mathews, D.C., and Tuner, D.H (1999). In: RNA Biochemistry and Biotechnology: Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide, NATO ASI Series, J. Barciszewski and B.F.C. Clark, eds. (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 11–43.

Received February 22, 2001; accepted June 15, 2001