Sequence-specific RNA Recognition by the Xenopus Y-box Proteins

AN ESSENTIAL ROLE FOR THE COLD SHOCK DOMAIN*

(Received for publication, July 7, 1995, and in revised form, September 5, 1995)

Philippe Bouvett§, Ken Matsumoto‡, and Alan P. Wolffe†

From the Laboratory of Molecular Embryology, NICHD, National Institute of Health, Bethesda, Maryland 20892-2710

The Xenopus Y-box protein FRGY2 has a role in the translational silencing of masked maternal mRNA. Here, we determine that FRGY2 will recognize specific RNA sequences. The evolutionarily conserved nucleic acid-binding cold shock domain is required for sequence-specific interactions with RNA. However, RNA binding by FRGY2 is facilitated by N- and C-terminal regions flanking the cold shock domain. The hydrophilic C-terminal tail domain of FRGY2 interacts with RNA independent of the cold shock domain but does not determine sequence specificity. Thus, both sequence-specific and nonspecific RNA recognition domains are contained within the FRGY2 protein.

The Xenopus laevis Y-box protein FRGY2 is a major component of ribonucleoprotein storage particles containing maternal mRNA within Xenopus oocytes (Murray et al., 1992; Deschamps et al., 1992; Tafuri and Wolfe, 1993a, 1993b). FRGY2 has an active role in facilitating the translational silencing or masking of maternal mRNA (Richter and Smith, 1984; Ranjan et al., 1993; Bouvet and Wolfe, 1994). How FRGY2 interacts with mRNA and how mRNA packaged by the FRGY2 protein is rendered accessible to the translational machinery is not understood.

The Y-box proteins contain a nucleic acid-binding domain conserved between prokaryotic and eukaryotic organisms (reviewed by Wolfe, 1994a, 1994b). The prokaryotic Y-box proteins regulate the cold shock response (Goldstein et al., 1990; La Teana et al., 1991; Jones et al., 1992). These proteins contain a single nucleic acid-binding structure known as the cold shock domain (CSD)1 (Wistow, 1990). The cold shock domain is a five-stranded β-barrel containing a well-characterized RNA binding motif RNP-1 (Schindelin et al., 1993; Schnuckel et al., 1993; Landsman, 1992; Burd and Dreyfuss, 1994a, 1994b). The cold shock domain will bind selectively to duplex (La Teana et al., 1991; Jones et al., 1992) and single-stranded DNA (Schindelin et al., 1993; Schnuckel et al., 1993). However, an exact DNA binding specificity is yet to be determined.

The eukaryotic Y-box proteins have been reported to bind with some selectivity to a wide variety of duplex and single-stranded nucleic acids including the Y-box (CTGTATTGCCCAA) (Marello et al., 1992; Wolfe et al., 1992; Murray, 1994) (reviewed by Wolfe, 1994a); however, the exact nucleic acid binding specificity was not determined. The eukaryotic Y-box proteins recognize RNA through the conserved CSD (Kolluri et al., 1992; Tafuri and Wolfe, 1992). The recognition of RNA has been proposed to be either predominantly via the CSD (Ladomy and Sommerville, 1994) or through interactions with islands of basic/aromatic amino acids in the C-terminal domain of those Y-box proteins (Murray, 1994).

In this work we established that both the oocyte-specific Y-box protein FRGY2 and a somatic homolog FRGY1 can have highly specific interactions with RNA. We determine that the cold shock domain is required for specific recognition of RNA. The C-terminal domain of FRGY2 has nonspecific interactions with RNA that may account for the more efficient association of wild type FRGY2 with RNA in comparison with the cold shock domain alone. We discuss the significance of these observations for the potential roles of FRGY2 in transcriptional and translational control.

MATERIALS AND METHODS

Proteins—To obtain the glutathione S-transferase (GST)-fusion constructs for the full-length (wild-type) or deletion mutants of FRGY2 protein, FRGY2 cDNA (see Tafuri and Wolfe (1990)) was generated using PCR amplification by Vent DNA polymerase (New England Biolabs) with oligonucleotide primer sets of either K9 (5′-CCGGCAATTC-CATGAGTGGGAGGCCAAGCCG-3′, for N terminus of FRGY2), K20 (5′-CGCAATTCCCAAAGAAGGTCGCCCACCTCAA-3′, based on amino acids (aa) 35–41 of FRGY2), or K7 (5′-CCGGCAATTCCGGAGGCTGTCAGATTAAAG-3′, based on aa 112–118) and either K19 (5′-CCGGCTG-GAGAAGCTTCTAGGGCTGCATATTGCCGC-3′, based on aa 111–105) or K8 (5′-CCGGCTGCAGAAGCTTCTAGGGCTGCATATTGCCGC-3′, for C terminus of FRGY2). The PCR products were digested with EcoRI and XhoI and isolated from an agarose gel. The fragments were then cloned into EcoRI– and XhoI-digested pGEX-4T-3 vector (Pharmacia Biotech Inc.).

Two steps of PCR amplification were performed to generate the constructs for point mutants of FRGY2 protein. Primers K12 (5′-GGT-GATACTCTCGGGCGCTTGCGAAAGTTAAAACC-3′; mutations corresponding to aa 55 and 57 of FRGY2 are underlined) and K13 (5′-AAGGGCCCGGACGTATCATTACACGAAATGACACAAAA-3′) were synthesized to obtain constructs for PM1 and CP1 proteins, and K14 (5′-GTGATAAA(G/A)CCGAAACCTTGGTCAACGGTTAAAACC-3′) and K15 (5′-AACCGGTTCGGAATCATTACCAACGAAATGACACAAAA-3′) were for PM2, CP2, and CP3 proteins. FRGY2 DNA was amplified by PCR with primer sets of K9/K12 and K13/K8 for PM1 or K9/K14 and K15/K8 for PM2. PCR products were isolated and used for secondary PCR amplification with the primer set of K9/K12, K9/K14, and K13/K8. The products from secondary PCR were used as the templates of PCR to generate constructs for CP1, CP2, and CP3 with the primer set of K9/K12. The final products were cloned into pGEX-4T-3 vector as described above. All constructs were completely sequenced to exclude the possible introduction of unintended fortuitous mutations during amplification and cloning.

To obtain GST-fusion FRGY2 protein and its derivatives, BL21(DE3) pLysS was transformed with each pGEX construct. A 300–500-ml culture was induced by isopropyl-1-thio-D-galactopyranoside to synthe-
size GST-fusion proteins. C-terminal deletion mutants D4, CSD, CP1, CP2, and CP3 were purified from bacterial lysate using glutathione-Sepharose 4B according to the method suggested by the manufacturer (Pharmacia). Since wild type, D5, D6, PM1, and PM2 proteins did not bind to glutathione-Sepharose efficiently, they were purified from the pelleted inclusion bodies of *Escherichia coli* lysate through SP-Sepharose chromatography as described (Tafuri and Wolffe, 1992), except that the step of precipitation with ammonium sulfate was omitted and the proteins were eluted from SP-Sepharose with phosphate buffer containing 0.75M NaCl after washing the column with the buffer containing 0.5M NaCl.

The full-length FRGY1 and FRGY2 proteins with T7 gene 10 leader peptides (T7-FRGY2) were expressed in *E. coli* BL21(DE3) transformed with the pET constructs (Tafuri and Wolffe, 1992). The Y-box proteins were purified from the pelleted inclusion bodies of *E. coli* lysate through an SP-Sepharose column as described above.

Random RNA and Selection of Preferred Sequences for FRGY2 Binding (Selex)—RNA was prepared as described previously (Tsai et al., 1991). Briefly, double-stranded transcription template was prepared by subjecting 5 ng of linear N25 oligodeoxynucleotide (Tsai et al., 1991) to 35 cycles in an Ericomp temperature cycler (1 min at 94°C, 1 min at 50°C, 2 min at 72°C) in the following buffer: 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.01% gelatin, 0.1 mg of primer T7Univ, 0.1 mg of primer RevUniv, 200 μM dNTP, and 2.5 units of Taq DNA polymerase. Any tandem products were reduced to monomers by cutting with *Bam*HI. The DNA template was then transcribed using T7 RNA polymerase.

RNA was selected from this random pool by coimmunoprecipitation with the FRGY1 and FRGY2 fusion proteins (Tsai et al., 1991). 1–5 μg of full-length T7 gene 10 leader peptides (T7-FRGY2) were expressed in *E. coli* BL21(DE3) transformed with the pET constructs (Tafuri and Wolffe, 1992). The Y-box proteins were purified from the pelleted inclusion bodies of *E. coli* lysate through an SP-Sepharose column as described above.

Random RNA and Selection of Preferred Sequences for FRGY2 Binding (Selex)—RNA was prepared as described previously (Tsai et al., 1991). Briefly, double-stranded transcription template was prepared by subjecting 5 ng of linear N25 oligodeoxynucleotide (Tsai et al., 1991) to 35 cycles in an Ericomp temperature cycler (1 min at 94°C, 1 min at 50°C, 2 min at 72°C) in the following buffer: 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 0.01% gelatin, 0.1 μg of primer T7Univ, 0.1 μg of primer RevUniv, 200 μM dNTP, and 2.5 units of Taq DNA polymerase. Any tandem products were reduced to monomers by cutting with *Bam*HI. The DNA template was then transcribed using T7 RNA polymerase.

RNA was selected from this random pool by coimmunoprecipitation with the FRGY1 and FRGY2 fusion proteins (Tsai et al., 1993). 1–5 μg of full-length T7 gene 10 fusion proteins was bound to 4 mg of protein A beads using the antibody against the gene 10 epitope (Novagen). After three washes with NT2 buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.05% Nonidet P-40, 1 mM MgCl2), the protein and 200–500 ng of RNA were incubated for 7 min in 100 μl of reaction mixture (20 mM KCl, 150 mM NaCl, 50 mM Tris, pH 7.5, 0.08% Nonidet P-40, 2.5% polyvinyl alcohol, 1 mM MgCl2, 1 mM EDTA, 50 μg/ml poly(A), 2 μg/ml vanadyl ribonucleoside complex, 0.5 mg/ml tRNA, 125 μg/ml bovine serum albumin, 1
approximately 0.4 mg of each protein was electrophoresed in 4–20% gradient polyacrylamide gel containing SDS. The gel was stained with Coomassie Brilliant Blue. D, GST-fusion FRGY2 proteins were tested for RNA binding activity (see “Materials and Methods”). The 32P-Labeled 2.14 wt RNA probe (95 nt) (lanes 2 and 12) was incubated with 0.1 μg (lanes 3, 5, 7, 9, 13, and 15) or 0.25 μg (lanes 4, 6, 8, 10, 14, 16, and 18) of GST-fusion FRGY2 proteins and subjected to gel retardation assay. The 2.14 wt RNA probe, which was heat-treated at 95 °C for 5 min and chilled on ice immediately, was electrophoresed in parallel (lanes 1 and 11). The positions of two RNA-protein complexes formed between GST-FRGY2 (WT) and the probe are indicated by dots.

**Fig. 2.** RNA binding of GST-fusion FRGY2 proteins. A and B, diagrams of FRGY2 protein and its derivatives. The cold shock domain (striped box), RNP-1 motif (filled box), and the basic (+) and acidic (−) amino acid clusters are indicated on the wild type FRGY2 protein. Deletion and point mutants are shown by open boxes. C, purification of FRGY2 proteins and GST-fusion derivatives used in this study. Approximately 0.4 μg each protein was electrophoresed in 4–20% gradient polyacrylamide gel containing SDS. The gel was stained with Coomassie Brilliant Blue.

Reverse transcription and reconstitution of the transcription template were as described previously (Tsai et al., 1991). RNA was reverse transcribed using 0.1 μg of primer RevUni and avian myeloblastosis virus reverse transcriptase (1 h at 42 °C using conditions recommended by the supplier). Following reverse transcription, the cDNA was resuspended in 10 μl of double distilled water. Three μl of the cDNA was subjected to 35 cycles of PCR under the conditions already noted. The new transcription template was then used to repeat the above transcription, comonomerprecipitation, and reverse transcription steps for four additional rounds. After the last PCR the cDNA was digested with BamHI and then subcloned in pGemBamHI. Then the cDNA were sequenced; 13 sequences for Y1 and Y2 contained a 6-nucleotide (nt) sequence that is very similar in all these cases (see Fig. 1). To introduce mutations into the FRGY2-binding site, three double-stranded oligonucleotides containing the sequences shown (5'-AACCGCGGTTATAGGTATG-3' or 5'-AAATCACATT-3') were used as the templates for new transcription template was then used to repeat the above transcription, comonomerprecipitation, and reverse transcription steps for four additional rounds. After the last PCR the cDNA was digested with BamHI and then subcloned in pGemBamHI. Plasmid DNA was sequenced, and one of the mutants, M3/M4–1, was used to generate the nonspecific RNA for competition used in this study.

Filter Binding Assay—Labeled in vitro transcribed RNA (0.05 pmol) was incubated in the presence of different amounts of GST-FRGY2 or T7 FRGY2 protein in a final volume of 100 μl of RNA binding buffer (20 mM Tris-HCl, pH 7.4, 4 mM MgCl2, 100 mM KCl, 0.1% Triton X-100, 1 mM dithiothreitol, 0.1 mg/ml (GST-FRGY2) or 0.5 mg/ml (T7FRGY2) yeast tRNA, 2 μg/ml vanadyl ribonucleoside complex, 20 μg/ml bovine serum albumin) for 20 min at room temperature. RNA-protein complexes were separated from unbound RNA by filtration through a nitrocellulose filter (BA85, Schleicher and Schuell). After three washes with 100 μl of 20 mM Tris-HCl, pH 7.4, 4 mM MgCl2, 100 mM KCl, the filter was dried and then exposed with x-ray film. The bound RNA was calculated by counting the radioactivity retained on the filter.

Dimethyl Sulfate Footprinting—RNA-protein complexes were formed as described above. Then dimethyl sulfate (1 μl) was added for 5 min at room temperature. Alkylation reactions are stopped by adding 50 μl of stop solution (1 μl Tris-HCl, pH 7.5, 1 μl β-mercaptoethanol, 0.1 mM EDTA). RNA was immediately precipitated with 2 volumes of 100% EtOH and 0.1 volume of sodium acetate. RNA pellets were resuspended in 0.3 μl sodium acetate, 0.5% SDS, 5 mM EDTA. After extraction twice with phenol/CHCl3, RNA was precipitated again. Alkylation sites were then analyzed by primer extension. RNA sequencing was also performed by introducing dideoxy-NTP in the primer extension reaction. Extended products are analyzed on an 8% polyacrylamide gel.

Gel Retardation Assay—32P-Labeled 2.14 wt RNA (95 nt) was prepared by transcribing the EcoRI- and HindIII-digested pGem construct that contains a FRGY2 recognition sequence (YRS) with T7 RNA polymerase. For the standard assay, 10 fmol of 32P-RNA was incubated with FRGY2 protein or its derivatives for 15 min on ice in a 10-μl RNA binding buffer described above. The mixture was electrophoresed in an 8% polyacrylamide gel (acrylamide:bis 29:1) containing 5% glycerol in a 10-μl RNA binding buffer at room temperature. The gel was dried and subjected to autoradiography.

To obtain shorter RNA probes, double-stranded oligonucleotides K22K23 (5'-AACCGCTGAGCTCATTAGGCGGCGGCCATTAAGAAT-3') or K5/K6 (5'-GATCTTAATACGACTCAGGATAAGG-3') were used as the templates for in vitro transcription with T7 RNA polymerase. Transcription with K22K23 and K5/K6 gave transcripts of 30 and 15 nt, respectively, and the probe are indicated by dots.
these were isolated from a polycrylamide gel. Non-labeled RNA for competitors were synthesized by transcribing the EcoRI and HindIII-digested pGEM construct containing a 2.14 wt site (see Table II) or the HindIII-digested pGEM construct containing a M3/M4–1 mutated site with T7 RNA polymerase.

RESULTS AND DISCUSSION

Selection of a 6-Nucleotide Consensus Sequence Recognized by FRGY1 and FRGY2—We used the methodology of Keene and colleagues (Tsai et al., 1991) to select specific RNA sequences bound by FRGY1 and FRGY2 from a degenerate pool of transcripts using immunoprecipitation of protein-bound RNA and PCR. Sequence analysis of multiple clones, each representing a distinct RNA species communoprecipitated with FRGY1 or FRGY2, revealed a favored recognition sequence (Fig. 1A). The consensus for FRGY1 and FRGY2 is very similar (Table I) and contains the hexanucleotide 5′-AACAUC-3′. We next wished to confirm the apparent specificity of RNA binding by FRGY1 and FRGY2 and performed RNA footprinting using one of the isolated RNA sequences (Table I, RNA 2.14 wt), which contains this consensus sequence. Footprinting with dimethyl sulfate of complexes of FRGY1 (Fig. 1B, lanes 2–4) and FRGY2 (Fig. 1B, lanes 9–11) with the RNA 2.14 wt revealed specific interactions with both Y-box proteins. We conclude that both FRGY1 and FRGY2 have a sequence-specific interaction with the hexanucleotide 5′-AACAUC-3′, a YRS.

Domains of FRGY2 Required for RNA Binding—We wished to determine which domains of FRGY2 were required for stable interaction with RNA. The FRGY2 protein can be divided into three putative domains (Fig. 2A); these include an N-terminal region, the CSD, and a C-terminal region containing peptide sequences that are alternately enriched in basic or acidic amino acids (Tafuri and Wolffe, 1990). The CSD contains a known RNA recognition motif, RNP-1 (Landsman, 1992). The CSD is the only highly conserved region of amino acid sequence between FRGY1 and FRGY2 (Tafuri and Wolffe, 1990) and therefore would a priori represent the domain most likely to confer sequence-specific recognition of RNA (Fig. 1). Using a variety of homopolymeric ribonucleotides, Ladomery and Sommerville (1994) determined that both the CSD and regions within the C-terminal tail of FRGY2 could interact with RNA. Murray (1994) suggested that the C-terminal tail was the primary site of interaction with RNA. These experiments made use either of chemically cleaved FRGY2 (Ladomery and Sommerville, 1994) or recombinant polypeptides (Murray, 1994) that contain partial fragments of the FRGY2 protein. We compared the properties of the intact C-terminal domain to the CSD and other deletion mutants of FRGY2 (Fig. 2A). The RNP-1 motif (GYGFI-) within the CSD (Fig. 2B) is a probable site of contact between FRGY2 and RNA (Burd and Dreyfuss, 1994a); therefore we introduced a series of point mutations into this domain (Fig. 2B). These mutations were based on previous work showing an essential role for conserved aromatic amino acids in the RNP-1 motif (Brennan and Platt, 1991; Merrill et al., 1988; Mayeda et al., 1994).

Full-length FRGY2 was purified either as a GST-fusion protein or as a T7 gene 10 fusion protein (Fig. 2C) (see “Materials and Methods”). The mutant proteins were purified as GST fusions (Fig. 2C). The C-terminal tail of FRGY2 is an extended hydrophilic structure extremely sensitive to proteolysis (Descamps et al., 1992). Thus, immunoblotting using polyclonal antibodies against FRGY2 reveals that the other proteins of smaller size than the full-length expressed protein are degradation products (data not shown). Gel retardation experiments (Fig. 2D) using a 95-nucleotide RNA probe containing the specific YRS show the GST-FRGY2 protein and the deletion mutants retaining the CSD together with either the N-terminal (D4) or C-terminal domain (D5) bound efficiently to the probe (Fig. 2D, lanes 3 and 4, lanes 15 and 16, and lanes 9 and 10, respectively). The CSD or the C-terminal tail domain alone (D6) bound less well to the probe (Fig. 2D, lanes 17 and 18 and lanes 13 and 14, respectively). Point mutations of both aromatic amino acids in the RNP-1 motif (PM1) reduced RNA binding (Fig. 2D, lanes 5 and 6), whereas a single point mutation (PM2) had relatively little effect on RNA binding within the context of the full-length protein (Fig. 2D, lanes 7 and 8). Control experiments (not shown) indicated that the GST moiety has no RNA binding activity in gel retardation (or filter binding) assay. We conclude that both the CSD and the C-terminal tail can contribute to the association of FRGY2 with RNA. Within the context of the full-length protein, the RNP-1 motif within the CSD has an important role. Both the N- and C-terminal do-

---

**Table II.** RNA binding by FRGY2

| Sequence       | Binding |
|----------------|---------|
| 2.14 wt        | AACATC  |
| M5/M6–7        | GACATC  |
| 6.3            | AACCTC  |
| 6.7            | AATGTG  |
| 6.10           | AACATC  |
| M5/M6–11       | AACTTA  |
| 6.2            | AATATC  |
| 6.11           | AACTAC  |
| M3/M4–2        | CACATC  |
| M3/M4–4        | CACAGG  |
| M5/M6–3        | GACATT  |
| M5/M6–6        | GACATA  |
| M3/M4–1        | AACTAG  |
| M3/M4–5        | AACTAG  |
| M5/M6–1        | AACATA  |

---

Fig. 3. RNA binding curves for FRGY2 and mutants. A filter binding assay (see “Materials and Methods”) was performed with GST-FRGY2 (○), PM1 (□), D4 (■), D6 (△), and the CSD (×). A, 32P-labeled RNA containing the consensus sequence (2.14 wt, AACAUC) was used as the probe. B, one of the point mutants in the consensus sequence (M3/M4–1, AACAUG, see Table II) was used as the probe. The data show the fraction of RNA bound as a percentage.
GST-fusion FRGY2 deletion mutants containing the CSD.

dation assay using the 2.14 wt RNA probe (Fig. 5) in parallel. The filter binding analysis of protein binding to the YRS consensus (Fig. 3A) reveals that GST-FRGY2 recognizes the 2.14 wt RNA probe with highest affinity \( K_d = 0.2 \times 10^{-7} \) M. The deletion mutants containing the CSD or tail domain alone (D6), the N-terminal domain including the CSD (D4), or point mutations in the RNP-1 motif all have reduced affinity \( K_d = 0.8-1 \times 10^{-7} \) M. The use of a nonspecific RNA sequence (M3/4.1) reduced the binding of GST-FRGY2 (to a \( K_d \) of 1.0 \( \times 10^{-7} \) M) (Fig. 3B). We conclude that the filter binding experiments confirm the general indications of affinity derived from gel shift analysis (Fig. 2).

We extended this analysis to the properties of the CSD in isolation. Using the gel retardation assay, we found that point mutations of the RNP-1 motif (Fig. 4A) dramatically reduced stable interaction with RNA probes (Fig. 4B). Quantitative measurements revealed a reduction in \( K_d \) from \( \sim 10^{-7} \) M for CSD to less than \( 10^{-6} \) M in the mutant proteins (not shown). A single point mutation of the RNP-1 motif has a more severe consequence for the CSD in isolation (Fig. 4B, lanes 10–12) than for the full-length FRGY2 protein (Fig. 2D, lanes 7–8). These differences might be due to the stabilizing influence of the C-terminal tail domain on FRGY2 interaction with RNA. Comparable effects are obtained when FRGY2 binds to DNA (Tafuri and Wolffe, 1992).

The association of FRGY2 with DNA is dependent on the length of the DNA probe used to assess binding (Tafuri and Wolffe, 1992). In our experiments we use RNA probes of 95 nt in length. Since earlier studies on selective RNA binding have generally made use of short 47-nt probes (Murray, 1994) or homopolymers of ill defined length (Ladomery and Sommerville, 1994), we explored whether RNA binding depends on the length of RNA probe. We find that little or no stable association of FRGY2 occurs with an RNA probe of 30 or 15 nt in length even though the YRS is present, whereas efficient binding occurs with a probe of 95 nt (Fig. 4C). Thus some of the reported discrepancies in nucleic acid binding selectivity may depend on the length of the probe and/or competitor in nucleic acid binding studies.

The CSD Confer Sequence-specific RNA Binding—Our experiments have defined an RNA sequence that is specifically recognized by the FRGY2 protein (Fig. 1, Table I). We have also determined that at least two RNA-binding domains exist within FRGY2 (Fig. 2). One of these is within the CSD (Figs. 2–4), and the other is within the C-terminal tail domain (Fig. 2). We next examined whether either of these domains conferred sequence-specific interactions with RNA. We initially determined what constituted a nonspecific RNA for FRGY2 by a combination of filter binding and gel retardation experiments in which the YRS consensus was systematically mutated (Fig.

**Fig. 4. Requirements for RNA binding of FRGY2.** A, diagram of GST-fusion FRGY2 deletion mutants containing the CSD. B, gel retardation assay using the 2.14 wt RNA probe (lane 1) was performed with 0.1 \( \mu \)g (lanes 2, 4, 7, 10, and 13), 0.25 \( \mu \)g (lanes 3, 5, 8, 11, and 14), and 0.625 \( \mu \)g (lanes 6, 9, 12, and 15) of GST-fusion FRGY2 proteins. C, binding of FRGY2 protein to RNA probes of various lengths. Gel retardation assays were performed with 10 fmol of 95 nt (lanes 2–5), 50 fmol of 30 nt (lanes 6–9), or 50 fmol of 15 nt (lanes 10–13) RNA probe containing a 2.14 wt binding site. The amount of T7-FRGY2 was 0 \( \mu \)g (lanes 2, 6, and 10), 0.04 \( \mu \)g (lanes 3, 7, and 11), 0.1 \( \mu \)g (lanes 4, 8, and 12), and 0.25 \( \mu \)g (lanes 5, 9, and 13). The positions of free probes and RNA-protein complexes are shown by asterisks and dots, respectively. MspI-digested pBR322 was electrophoresed in parallel (lane 1).
In these experiments we increased the excess of competitor mRNA in the binding reaction from that in the earlier filter binding experiments (Fig. 3) (see "Materials and Methods"). This was done in order to provide a clearer discrimination in the binding of FRGY2 to different RNA sequences. Our results (shown in Fig. 5 and summarized in Table II) reveal that every ribonucleotide in the YRS has an important role in FRGY2 binding to RNA. In these experiments we make use of T7-FRGY2 (see "Materials and Methods"). Our estimation of binding affinity for the YRS containing RNA probe (Fig. 5, \( K_d = 0.2 \times 10^{-7} \) M) is very similar to that for the GST-FRGY2 protein (Fig. 3A, \( K_d = 0.2 \times 10^{-7} \) M). Thus, independent expression systems yield comparable binding values, indicative of equivalent recovery of active protein. We chose a 100-nt RNA in which the sequence AACAUG (Table II, M3/M4-1) replaced the YRS as a nonspecific competitor (see also Fig. 3B).

Our gel retardation experiments with T7-FRGY2 reveal that two complexes assemble on the 95-nt YRS RNA as the excess of protein over probe is increased (Fig. 6A, lanes 1–5). Since identical bands are seen with nonspecific RNA that does not contain a YRS (Tafuri and Wolffe, 1993a; Tafuri et al., 1993), this may reflect both specific and nonspecific protein-RNA interactions. In vivo the FRGY2 protein interacts with mRNA such that one protein binds every 40 nucleotides (Darnbrough and Ford, 1981). We chose an excess of FRGY2 over the RNA probe that led to the accumulation of a single complex (Fig. 6A, lane 6) and competed this complex with an increasing excess of RNA competitors (Fig. 6B, lanes 7–12). Competition with the 95-nt RNA containing the YRS was much more effective than when this sequence was altered.

We next examined whether deletion mutants of FRGY2 containing either the CSD alone, the N terminus including the CSD, or the C-terminal portions of the protein would have specific interactions with RNA. The C-terminal tail domain of FRGY2 binds to the specific RNA probe, but this interaction is
Specific RNA Sequence Recognition by Y-box Proteins

competed equivalently by both specific and nonspecific RNA competitors (Fig. 6B, lanes 7–9). In contrast, the CSD alone (Fig. 6B, lanes 1–3) and the N terminus of FRGY2, which includes the CSD (Fig. 6B, lanes 4–6), bind to the specific RNA probe such that binding is selectively competed by the YRS sequence. This suggests that the CSD contributes to the specific recognition of RNA but requires the flanking N- and C-terminal domains of FRGY2 for a high affinity interaction (see Fig. 3). The CSD is the only conserved amino acid sequence in the N-terminal portion of FRGY1 and FRGY2 (Tafuri and Wolfe, 1990) leading to recognition of the same YRS consensus. Since RNA binding is weakened by mutation of the RNP-1 motif (Fig. 2D, lanes 5 and 6, Fig. 3), it appears that the RNP-1 motif may have a role in maintaining the structural integrity of the CSD necessary for discrimination between various RNA sequences.

Biological Implications—The Y-box proteins have the capacity to interact with RNA with sequence specificity (Figs. 1, 3, 5, and 6 and Tables I and II). The highly conserved cold shock domain is required for this sequence-specific recognition of RNA (Figs. 2–6). This domain differs in only four out of seventy-eight amino acids between FRGY1 and FRGY2, whereas these proteins are highly divergent at the N terminus and display an overall identity of 65% (Tafuri and Wolfe, 1990). Important, other metazoan Y-box proteins all retain the same cold shock domain sequences as FRGY1 or FRGY2 and will thus have the capacity to interact specifically with RNA. This might influence their capacity to regulate both transcriptional and translational events (Ranjan et al., 1993; Bouvet and Wolfe, 1994).

In the oocyte, FRGY2 and a highly related protein mRNP3 package mRNA (Deschamps et al., 1992; Murray et al., 1992; Tafuri and Wolfe, 1993a, 1993b). These proteins interact with mRNA with a stoichiometry such that one protein binds approximately every 40 nucleotides (Darnbrough and Ford, 1981; Tafuri and Wolffe, 1993a, 1993b). These proteins interact with the vast majority of cytoplasmic mRNA (Tafuri and Wolffe, 1993a, 1993b). The Y-box proteins might interact with specific nascent RNAs and communicate with the basal transcriptional machinery in a manner analogous with that of TAT-TAR (Berkhout et al., 1989). Although RNA binding by the prokaryotic Y-box proteins has not yet been examined, it is possible that sequence-selective RNA binding might contribute to regulation of the cold shock response (Goldstein et al., 1990). A requirement to maintain the capacity to interact with a variety of specific nucleic acids might contribute to the striking evolutionary conservation of the cold shock domain itself.

Acknowledgments—We thank Dr. J. Keene for instruction in Selex strategies. We thank Thuy Vo for manuscript preparation.

REFERENCES

Berkhout, B., Silverman, R. H., and Jhang, K.-T. (1989) Cell 59, 273–282
Bouvet, P., and Wolfe, A. P. (1994) Cell 77, 931–941
Brennan, C. A., and Platt, T. (1991) J. Biol. Chem. 266, 17296–17305
Burd, C. G., and Dreyfuss, G. (1994b) EMBO J. 13, 1197–1204
Darnbrough, D. H., and Ford, P. J. (1981) Eur. J. Biochem. 113, 415–424
Deschamps, S., Viel, A., Garrigos, M., Denis, H., and le Maire, M. (1992) J. Biol. Chem. 267, 13799–13802
Evdokimova, V. M., Wei, C.-L., Sitikov, A. S., Simonenko, P. N., Lazarev, O. A., Vlasilenko, K. S., Ustunov, V. A., Hershley, J. W. B., and Ovchinnikov, L. P. (1995) J. Biol. Chem. 270, 3186–3192
Goldshlager, N., Pollitt, N. S., and Inouye, M. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 283–287
Grundstein, M. (1990) Annu. Rev. Cell Biol. 6, 643–678
Jones, P., Krah, R., Tafuri, S. R., and Wolfe, A. P. (1992) J. Biol. Chem. 267, 5798–5802
Kaschani, F., Duvall, J. F., Dittmer, J., Mireskandari, A., Reid, R. L., Gitlin, S. D., and Brady, J. N. (1994) J. Virol. 68, 561–565
Kerr, D., Chang, C., Chen, N., Galli, G., Raj, S., Schwartz, B., and Khaliili, K. (1994) J. Virol. 68, 7637–7643
Kolluri, R., Torrey, T. A., and Kinniburgh, A. J. (1992) Nucleic Acids Res. 20, 111–116
La donoroy, M., and Somervillle, J. (1994) Nucleic Acids Res. 22, 5582–5589
Lansdorp, P. M. (1992) Nucleic Acids Res. 20, 2861–2864
La Teana, A., Bandi, A., Falconi, M., Sprino, R., Pon, C. L., and Gualerzi, C. O. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 10907–10911
MacDonald, G. H., Itoh-Lindstrom, Y., and Ting, J.-Y. (1995) J. Biol. Chem. 270, 3527–3533
Marello, K., La Rovere, J., and Somervillle, J. (1992) Nucleic Acids Res. 20, 5993–5997
Mayeda, A., Monroe, S. H., Caeceres, J. F., and Krainer, A. R. (1994) EMBO J. 13, 5483–5495
Merrill, B. M., Stone, K. L., Cobbiani, F., Wilson, S. H., and Williams, K. R. (1988) J. Biol. Chem. 263, 3307–3313
Murray, M. T. (1994) Biochemistry 33, 13910–13917
Murray, M. T., Schiller, D. L., and Franke, W. W. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11–15
Ranjan, M., Tafuri, S. R., and Wolfe, A. P. (1993) Genes & Dev. 7, 1725–1736
Richter, J. D., and Smith, L. D. (1984) Nature 309, 378–380
Schnidelin, H., Marahel, M. A., and Moxemann, U. (1993) Nature 364, 164–169
Schnuckel, A., Wiltscheck, R., Czsich, M., Herrier, M., Willmsky, G., Graumann, P., Marahel, M. A., and Holak, T. A. (1993) Nature 364, 169–171
Tafuri, S. R., and Wolfe, A. P. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 9028–9032
Tafuri, S. R., and Wolfe, A. P. (1992) New Biol. 4, 349–359
Tafuri, S. R., and Wolfe, A. P. (1993a) J. Biol. Chem. 268, 24255–24261
Tafuri, S. R., and Wolfe, A. P. (1993b) Trends Cell Biol. 3, 94–98
Tafuri, S. R., Familiar, M., and Wolfe, A. P. (1993) J. Biol. Chem. 268, 12213–12220
Ting, J., Painter, A., Zalednik-Le, N. J., MacDonald, G., Moore, T. M., Brown, A., and Schwartz, B. D. (1994) EMBO J. 13, 1605–1611
Tsai, D. E., Harper, D. S., and Keene, J. D. (1991) Nucleic Acids Res. 19, 4931–4936
Wistow, G. (1990) Nature 344, 823–824
Wolffe, A. P. (1992) Chromatin, Structure and Function, Academic Press, San Diego, CA
Wolffe, A. P. (1994a) BioEssays 16, 245–251
Wolffe, A. P. (1994b) Trends Biochem. Sci. 19, 240–244
Wolffe, A. P., Tafuri, S., Ranjan, M., and Familiar, M. (1992) New Biol. 4, 290–298
Sequence-specific RNA Recognition by the *Xenopus* Y-box Proteins: AN ESSENTIAL ROLE FOR THE COLD SHOCK DOMAIN
Bouvet Philippe, Ken Matsumoto and Alan P. Wolffe

*J. Biol. Chem.* 1995, **270**:28297-28303.
doi: 10.1074/jbc.270.47.28297

Access the most updated version of this article at [http://www.jbc.org/content/270/47/28297](http://www.jbc.org/content/270/47/28297)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 38 references, 16 of which can be accessed free at [http://www.jbc.org/content/270/47/28297.full.html#ref-list-1](http://www.jbc.org/content/270/47/28297.full.html#ref-list-1)