Binding of Y-box proteins to RNA: involvement of different protein domains

Michael Ladomery and John Sommerville*
School of Biological and Medical Sciences, University of St Andrews, St Andrews, Fife KY16 9TS, UK

Received September 27, 1994; Revised and Accepted November 18, 1994

ABSTRACT

Eukaryotic Y-box proteins are reported to interact with a wide variety of nucleic acid structures to act as transcription factors and mRNA masking proteins. The modular structure of Y-box proteins includes a highly conserved N-terminal cold-shock domain (CSD, equivalent to the bacterial cold-shock proteins) plus four basic C-terminal domains containing arginine clusters and aromatic residues. In addition, the basic domains are separated by acidic regions which contain several potential sites for serine/threonine phosphorylation. The interaction of Y-box proteins, isolated from Xenopus oocytes (FRGY2 type), with RNA molecules has been studied by UV crosslinking and protein fragmentation. We have identified two distinct binding activities. The CSD interacts preferentially with the polypurines poly(A,G) and poly(G) but not poly(A), this activity being sensitive to 5 mM MgCl₂ but not to 5 mM spermidine. In the presence of 1 mM MgCl₂ or 1 mM spermidine, the basic domains interact preferentially with poly(C,U), this activity being sensitive to 0.5 M NaCl. Binding of the basic domains is also sensitive to low concentrations of heparin. The basic domains can be crosslinked individually to labelled RNA. These results are discussed with reference to the various specificities noted in the binding of Y-box proteins to RNA and DNA.

INTRODUCTION

The Y-box proteins are the most evolutionarily conserved nucleic acid-binding proteins yet described, found in bacteria, plants and animals (1,2). The eukaryotic Y-box proteins were originally identified through their ability to interact with DNA containing a reverse CCAAT box, the Y-box sequence CTGATTGGCCAA (3). This sequence is found in a variety of promoter regions, including those of the MHC class II genes (3,4) and genes encoding germ cell-specific functions (1) and in these contexts the Y-box proteins are considered to act as regulators of transcription. However, a range of Y-box proteins were subsequently characterized through their selective interaction with promoter sequences containing pyrimidine-rich single-stranded DNA (5—7), apurinic DNA (8,9) and even purine-rich single-stranded DNA (10). Furthermore, the most abundantly expressed Y-box proteins, the FRGY2 class from Xenopus oocytes (11) and the MSY1 class from mouse spermatocytes, are found bound to mRNA (12,13). Thus a universal family of highly conserved proteins has been reported to recognize a diversity of nucleic acid structures. Either the remarkable versatility in nucleic acid recognition is a special property of all Y-box proteins and the in vitro conditions for their diverse binding properties have not yet been properly defined, or else individual Y-box proteins have evolved different specificities. The Y-box proteins consist of a modular series of domains, each of which has the potential to bind nucleic acids. Therefore the presence of multiple binding domains and variations in their arrangement might explain the different binding specificities of different Y-box proteins.

The most highly conserved feature of eukaryotic Y-box proteins is the presence of a 70 amino acid domain which is 43% homologous to the cold-shock protein, CSPB, of E. coli (14). The structure of a corresponding protein, CSPB, from B. subtilis has been solved by crystallographic (15) and NMR (16) analysis and consists of an antiparallel five-stranded β-barrel. More recently, this structure has been confirmed for CSPB (17,18). The three N-terminal β-strands present a face with exposed aromatic and basic side chains which could interact with nucleic acids. This cold-shock domain (CSD) is sufficient to bind Y-box sequences in vitro and to drive the expression of bacterial cold-shock-inducible genes (19,20). Apart from the mammalian protein unr (encoded by an ORF located upstream of n-ras), which consists of five tandem repeats of a CSD (21), eukaryotic Y-box proteins contain a single CSD plus a series of C-terminal charged domains. FRGY2 is quite typical in containing four basic domains, rich in arginine and aromatic residues, separated by four domains rich in acidic residues and potential sites for serine/threonine phosphorylation (11,22,24). Arginine-rich domains are a common feature of many RNA-binding proteins (25) and their multiple presence in Y-box proteins increases the potential for nucleic acid binding.

In Xenopus oocytes, Y-box proteins are expressed at high concentration (~0.1 μg/oocyte) as an equimolar pair (here FRGY2α = pp60 and FRGY2β = pp56; cf 11). They can be readily isolated from native mRNP particles (26,27) and by

*To whom correspondence should be addressed
simple heat treatment of oocyte homogenates (28). The particular advantage of working with native rather than with recombinant forms is that they are isolated in a state of interaction with nucleic acid. Since phosphorylation of FRGY2 is relevant, and possibly crucial, for binding to mRNA (26,29,30) and until phosphorylation sites are pinpointed and manipulated in vitro, proteins properly modified in vivo are to be preferred.

In this report we examine further the RNA binding preferences of the FRGY2 proteins and discriminate between different binding domains by UV crosslinking and protein fragmentation.

**MATERIALS AND METHODS**

**Isolation of Y-box proteins**

Poly(A)+ mRNP was isolated from the previtellogenic ovary of *Xenopus laevis* as described previously (30). To select the Y-box proteins, the eluate from oligo(dT) cellulose (Pharmacia) was adjusted to 20 mM NaCl, 20 mM Tris–HCl, pH 7.5 (HTB), heated to 80°C for 10 min, cooled at room temperature for 5 min, chilled on ice for 5 min and centrifuged at 10,000 r.p.m. for 10 min (28). Under these conditions the Y-box proteins remain in the supernatant and can be recovered quantitatively as a complex with mRNA. Multiple cycles of heat treatment/centrifugation improve the purity of the preparation. All other mRNP proteins are denatured and can be recovered from the pellets. Riboprobes added to mRNP prior to heat treatment are recovered complexed with the Y-box proteins.

To separate the Y-box proteins from mRNA, the heat-treatment supernatant is subjected to a further cycle of heat treatment in the presence of heparin—Sepharose CL 6B (Pharmacia). About 100 mg of RNP and 0.5 ml of resin was suspended in 1 ml of HTB and heated to 80°C. After cooling, with continuous shaking, the slurry was pipetted into a column and, after settling, was rinsed with HTB. The RNA was eluted either with 8 M urea or with 5 mM MgCl₂. 8 M urea also washed off contaminating proteins. Finally the Y-box proteins were eluted with 1 M NaCl, dispensed into aliquots and stored at -70°C. The purity of preparations was checked by SDS–PAGE, the Y-box proteins having apparent molecular weights of 60 and 56 kDa.

**Synthesis of riboprobes**

The RNA used for binding studies corresponds to the 3' end (280 nucleotides) of an mRNA encoding an oocyte-specific β-tubulin. Over 90% of this mRNA is packaged as non-polysomal mRNP particles and is representative of mRNAs associated with Y-box proteins in *Xenopus* oocytes (N. Clark, M. Ladomery, and J. Sommerville, unpublished). The sequence is: GAAUUCACUG AGGCGAGAG CAACAUGAAC GACCUGGGU CUGAGUCCA ACAGUACCAG GAUCCCGCG CUGAGGAGGA GGGAGAGUU AAGAGGAGA AAAUGCC-UAA AGCUCCUUUA CACUUGUAAA UUAUCAUCU CUAUCUGUC CUGGUUUCA UUAUUGUGU ACCU-GCAU UGUCCUCUC CAGUUCUAU GUUACCAGU GUUACAGAAGC UUGCUUCAU UAAAAGCAGAU-UUUUCAUCUG AAAAAAAAAAAAAAAAAAAA.

The translation stop codon UAA, the polyadenylation motif AUUAAA and a potential cytoplasmic polyadenylation element UUUUCAU are underlined.

The RNA was synthesised by run-off transcription from a cDNA subclone in pBlueScript (Stratagene) in the presence of [α-32P]CTP as described previously (30). Double-stranded RNA was formed by annealing equimolar amounts of labelled sense strands and unlabelled antisense strands. Remaining single strands were digested with ribonuclease and the double-stranded probe was purified by phenol extraction and ethanol precipitation.

Various ribopolymers were dissolved at 10 mg/ml in distilled water (stock). The heteropolymers poly(C,U) and poly(A,G) were dephosphorylated and 5' ends were end-labelled using [γ-32P]ATP (3,000 Ci/mmol) and polynucleotide kinase as recommended on the enzyme data sheet (Pharmacia).

**Protein—riboprobe binding**

0.1 µg of riboprobe was added to 2 µg of Y-box proteins (or RNP) and binding was accomplished in a final volume of 50 µl by either: (i) dialysis from 1 M NaCl into HTB; or (ii) dilution from 20 to 2 mM MgCl₂ with HTB; or (iii) heat treatment as described above. Protein—riboprobe interaction was challenged by adding heparin, salts, urea or polynucleotides either before or after the binding step. Conditions for the crosslinking of RNA to proteins using UV light were as described previously (30). For quantitative assessment of crosslinking efficiency, 50 µl samples were adjusted to 1% SDS, 0.15 M NaCl, 5 mM EDTA, 50 mM Tris–HCl, pH 7.5, and extracted with an equal volume of phenol—chloroform (1:1 mixture). After phase separation, the radioactivity in aqueous and organic phases was measured and the percentage of initial radioactivity extracted into the organic phase was calculated. For identification of proteins and proteolytic fragments crosslinked to riboprobe, samples were digested with

![Figure 1](https://example.com/figure1.png)
These results charged heparin, through complexes formed released interaction with RNA. Subsequently released with are Sepharose (arginine hybridizing indicating crosslinked, sequence, aspartate-proline FRGY2). Methods). The binding of riboprobe to gel transfers was performed as follows. Transfers were washed twice for 10 min in 6 M urea in binding buffer (50 mM NaCl, 0.5% Tween-20, 10 mM Tris–HCl, pH 7.5). Riboprobe (0.5–1.0 μg) was added in 5–20 ml of binding buffer, then incubated for 30–60 min. Finally the transfers were washed twice for 10 min in binding buffer, either with 50 mM NaCl (low salt wash) or with 500 mM NaCl (high salt wash). The transfers were then set up for autoradiography.

Band-shift assays were used to assess the binding of FRGY2 proteins to end-labelled ribopolymers. Complexes, formed in the presence or absence of MgCl2 and competitors, were electrophoresed, without crosslinking, through 1.5% agarose in 20 mM Tris–HCl, pH 7.5. The gels were fixed in 10% acetic acid, dried and set up for autoradiography.

**Chemical proteolysis**

Aspartate–proline bonds were hydrolysed by incubating crosslinked FRGY2–riboprobe complexes (made using 10 μg of protein and 0.5 μg of riboprobe in a final volume of 50 μl) in 70% formic acid at 40°C for 24 h (31).

Cleavage at asparagine–glycine residues was achieved by incubating FRGY2 proteins in 2% hydroxylamine at 20°C for 4 h (32). The mixture was maintained at pH 9 with NaOH.

Chemically cleaved proteins were extensively dialysed against HTB before use in RNA binding.

**RESULTS**

**FRGY2 proteins can bind RNA and heparin simultaneously**

The Y-box proteins FRGY2a (pp60) and FRGY2b (pp56) can be purified from *Xenopus* oocyte mRNP particles in three steps: (i) selection of poly(A)+ RNA by affinity binding to oligo(dt) cellulose (30); (ii) heat treatment of the poly(A)+ RNA at 80°C followed by chilling and centrifugation (28); (iii) binding of the heat-treatment supernatant to heparin–Sepharose followed by salt elution of the Y-box proteins (described below). The RNA used in the binding studies described here represents 280 nucleotides of the 3′ end of an oocyte-specific β-tubulin mRNA (see Materials and Methods). Preliminary studies (not shown) confirmed that this sequence, but not its double-stranded form made by hybridizing sense and antisense strands, was highly efficient in binding FRGY2 proteins.

Complexes of FRGY2 proteins and RNA, formed in vitro but not crosslinked, do not bind to heparin–Sepharose (Fig. 1A), indicating that the arginine clusters in the protein tail domains are not accessible for binding to heparin and are neutralized through protein–RNA and/or protein–protein interaction. However, complexes formed in the presence of heparin–Sepharose are bound to the resin, the RNA component being released with 8 M urea or 5 mM MgCl2, and the proteins being subsequently released with 1 M NaCl (Fig. 1B and C; see also 5,7). These results indicate that regions of positive charge (arginine clusters) in the proteins interact with the negatively charged heparin, but not at the expense of all of the protein interaction with RNA. Thus, it is possible that recognition of RNA can be achieved independently of the basic charge domains, for instance through the CSD (2).

**RNA binding is blocked by a combination of heparin and Mg2+**

Points of contact of Y-box proteins with RNA can be fixed by crosslinking with UV irradiation (29,30). In principle, multiple RNA-binding sites within a single protein can be discriminated by differential blocking or by fragmenting the protein.

Contacts between FRGY2 proteins and RNA can be established in the presence of 10 μg/ml heparin, 3 mM MgCl2, 0.5 M NaCl or 4 M urea, the percentage of riboprobe radioactivity crosslinked to protein under these conditions being similar to the control value (Fig. 2A). However, certain combinations of these agents, namely heparin plus MgCl2, heparin plus urea and NaCl plus MgCl2, almost completely prevent crosslinking (Fig. 2A). Even RNA–protein complexes formed first and then treated with heparin/MgCl2 or heparin/urea before crosslinking are

![Figure 2](image-url)

**Figure 2. Effect of different agents on the UV crosslinking of riboprobe to FRGY2 proteins.** (A) The binding buffer (20 mM NaCl, 20 mM Tris–HCl, pH 7.5) was adjusted with the agents shown prior to the binding of proteins to RNA, UV irradiation and phenol extraction. (B) Confirmation of crosslinking to FRGY2a and FRGY2b. In this experiment, the buffer was adjusted with the agents shown after the binding reaction. Then the complexes were UV irradiated, digested with ribonuclease and analysed by SDS–PAGE/autoradiography. Note that in (A), a minimum of one crosslinking event per protein–riboprobe complex is sufficient to be recorded as maximum binding, whereas in (B), the intensity of labelling is proportional to the number of crosslinking events.
susceptible to dissociation (Fig. 2B). That heparin and urea should cooperate to prevent binding and to disrupt complexes is hardly surprising, since protein conformation and RNA recognition are largely dependent upon charge interaction and hydrogen bonding. However, the combined effect of low concentrations of heparin and Mg\(^{2+}\) is more interesting and potentially useful.

**Heparin and Mg\(^{2+}\) block different binding domains**

As mentioned earlier, the effect of heparin is most likely to block the positively charged tail domains of the Y-box proteins, leaving the CSD for interaction with nucleic acid (in the absence of Mg\(^{2+}\)). The additional effect of Mg\(^{2+}\) would therefore operate mainly through the CSD. That this is, indeed, the case is demonstrated by disrupting the CSD. In all vertebrate Y-box proteins whose sequence has been determined, there exists an unique NG (asparagine—glycine) site at the beginning of the second \(\beta\)-strand of the \(\beta\)-barrel. This site (…RNGYGFINR…) shows homology with the RNP-I site of the RRM-containing family of RNA-binding proteins (11,25). Cleavage of FRGY2 proteins at the NG site (see Fig. 5A) with hydroxylamine (HA) results in a protein preparation which still crosslinks to RNA in the presence or absence of 3 mM MgCl\(_2\); this binding showing the characteristics of heparin inhibition similar to those obtained with the native proteins (Fig. 3A). These results would be explained by a Mg\(^{2+}\)-induced block operating at, or near to, the CSD and a heparin block imposed by binding to the basic tail domains. That the Mg\(^{2+}\)-induced block applies to proteins with an intact CSD, but not to the HA-cleaved proteins, is seen in Fig. 3B. It is also seen that heparin inhibition of tail domain binding (in the HA-cleaved proteins) is activated by low concentrations of MgCl\(_2\) (<1 mM). One further consequence of disrupting the CSD is that the stability of interaction with the riboprobe is substantially reduced. Thus 0.5 M NaCl appears to be sufficient to prevent binding by the tail domains of the HA-cleaved proteins, yet has little effect on crosslinking to the intact proteins (Fig. 3C).

![Figure 3](image_url)

**Figure 3.** Effects of disrupting the CSD on the ability of FRGY2 proteins to bind riboprobe in the presence of heparin, MgCl\(_2\), and NaCl. Intact FRGY2 (Native) and proteins cleaved with HA were crosslinked to riboprobe in the conditions indicated. (A) Heparin-sensitive binding activity in the presence and absence of MgCl\(_2\) (3 mM). (B) Mg\(^{2+}\)-sensitive binding activity in the presence and absence of heparin (10 \(\mu\)g/ml). (C) Na\(^{+}\)-sensitive binding activity.

![Figure 4](image_url)

**Figure 4.** Effects of Mg\(^{2+}\) on the interaction of FRGY2 proteins with ribopolymers. (A) The Mg\(^{2+}\)-induced switch in binding specificity. The FRGY2−riboprobe interaction was challenged with a 100-fold excess over riboprobe of poly(C,U), poly(A) or poly(A,G) at the concentrations of MgCl\(_2\) shown or in the presence of 1 mM EDTA. (B) Band-shift assay showing direct binding of radiolabelled ribopolymers. The poly(A,G) probe (tracks 1−4) and the poly(C,U) probe (tracks 5−8) were bound to FRGY2 proteins in the absence (tracks 2−4 and 6) and presence (tracks 1, 5, 7 and 8) of 3 mM MgCl\(_2\) and in the presence of 100-fold excess of unlabelled poly(A) (tracks 1, 2, 5 and 6), poly(A,G), (tracks 3 and 8) and poly(C,U) (tracks 4 and 7). The positions of unbound probe (P), protein−RNA complexes (C) and larger aggregates (arrow) are indicated. Confirmation that protein−RNA complexes have been formed is given by the relative crosslinking values obtained from the corresponding reactions.
Sequence recognition is influenced by Mg$^{2+}$

We reported previously (30) that the FRGY2 proteins show a marked binding preference for polypurrimidines, the heteropolymer poly(C,U) being the best competitor of riboprobe binding tested. These experiments were performed in the presence of Mg$^{2+}$, diluted from 20 mM (to destabilize native mRNP complexes) to 2 mM MgCl$_2$ (to allow binding of the Y-box proteins to riboprobe). As shown here, this binding preference can now be largely attributed to the tail domains. The results (Fig. 4A) confirm that competition by poly(C,U) is actually Mg$^{2+}$-dependent. Of a range of polymers used to study competition in the absence of Mg$^{2+}$, by far the best competitors were the polypurines, poly(A,G) and poly(G), but as with Mg$^{2+}$-dependent binding (30), CSD binding showed no competition by poly(A) (Fig. 4A). The polypurine-binding activity of the CSD was progressively inhibited by 1–3 mM MgCl$_2$. In effect, a switch in binding affinity from poly(A,G) to poly(C,U) can be achieved simply by increasing the concentration of MgCl$_2$.

That FRGY2 proteins can interact directly with poly(A,G) or poly(C,U) depending on the binding conditions is shown by band-shift assays using labelled polymers (Fig. 4B). In the presence of 3 mM MgCl$_2$, poly(A,G) is not bound, whereas poly(C,U) is retarded as RNA–protein complexes (C). In the absence of MgCl$_2$, poly(A,G) forms complexes, whereas poly(C,U) is not bound. The binding reactions are unaffected by adding excess amounts of unlabelled poly(A), but are competed by adding the same polymer as the bound probe (Fig. 4B). Addition of excess poly(C,U) in the absence of Mg$^{2+}$ and excess poly(A,G) in the presence of Mg$^{2+}$ improves binding efficiency to the poly(A,G) and poly(C,U) probes, respectively. One further consequence of adding excess poly(A,G) to the reaction in which the tail domains bind the poly(C,U) probe is that the complexes are driven into larger aggregates (arrow in Fig. 4B). The molecular basis for this increased extent of interaction is not obvious.

**Fragmentation of FRGY2 proteins reveals multiple binding domains**

As already discussed, the CSD of the FRGY2 proteins can be disrupted with HA cleavage at a unique NG site (Fig. 5A). In addition, hydrolysis of proteins with formic acid (FA) results in preferential cleavage at aspartate–proline (DP) residues (31). A single DP site exists in FRGY2b (Fig. 5A), such that FA treatment results in an N-terminal fragment containing the CSD plus two basic domains and a C-terminal fragment containing two basic domains. Production of these fragments from isolated FRGY2b has been confirmed by immunostaining (not shown). In FRGY2a there are three further DP sites (Fig. 5A), resulting in major FA-sensitive cleavage products.

---

**Figure 5.** RNA binding to chemically cleaved FRGY2 proteins. (A) Diagram of FRGY2a/b showing their linear structure, consisting of: the N-terminal region (N); the CSD (β1–β5); the acidic domains (A1–A4); the basic domains (B1–B4). The position of the HA-sensitive NG site and the positions of the FA-sensitive DP sites are indicated by arrows. Also shown are the potential sites of phosphorylation by the RNP-bound casein kinase II (asterisks). (B) Cleaved fragments retain RNA-binding activity. FRGY2 proteins (tracks 1 and 1') and fragments produced by HA (tracks 2 and 2') and FA (tracks 3 and 3') treatments were separated by SDS–PAGE, transferred to nitrocellulose and either immunostained using anti-FRGY2 (tracks 1–3) or incubated with riboprobe to produce the autoradiograph shown (tracks 1'–3'). Positions of the major fragments C-terminal to the NG site (∆N31) and N-terminal to the DP sites in A2 (∆A2–A4) are indicated by arrows on the immunoblot. (C) Points of crosslinking (contact) are established throughout much of the length of the protein. FRGY2–riboprobe complexes, formed under different conditions, were crosslinked, digested with ribonuclease, cleaved with FA and analysed by SDS–PAGE/autoradiography. Complexes were formed from 10 μg of protein and 0.5 μg of riboprobe in binding buffer with no addition (tracks 1 and 2) or with addition of: 3 mM MgCl$_2$ (track 3); 10 μg of heparin (track 4); MgCl$_2$ plus heparin (track 5); 20 μg of poly(C,U) in the absence (track 6) or presence (track 7) of MgCl$_2$; 20 μg of poly(A,G) in the absence of MgCl$_2$ (track 8). Indicated by arrows are the positions of: the intact proteins (FRGY2); the N-terminal fragments (∆A2–A4); the C-terminal fragment from FRGY2b (B3 + B4); the C-terminal fragments from FRGY2a (B3 and B4); (D) Individual basic tail domains can bind RNA in the absence of a functional CSD. HA-cleaved proteins were bound to riboprobe and then crosslinked and treated with FA as described above. Complexes were formed in the absence (tracks 1 and 2) or presence (tracks 3–5) of 3 mM MgCl$_2$ and with the addition of 2 μg heparin (track 4) or 20 μg of poly(C,U) (track 5). Fragments could be aligned with those produced by FA treatment alone and run on the same gel (track 1). Identity of fragments is shown for (C).
in a similar N-terminal fragment plus two small fragments each containing a single basic domain.

The major HA-cleaved fragments (ΔN71) and FA-cleaved fragments (ΔA2-A4) can be clearly seen in the immunoblot shown (Fig. 5B, tracks 2 and 3). That these fragments retain RNA-binding activity is seen in the duplicated blot, which has been incubated with labelled riboprobe (Fig. 5B, tracks 2’ and 3’). In this particular assay, binding was accomplished in the absence of Mg2+, therefore labelling of fragments with disrupted CSDs (ΔN71) is poor in comparison with fragments containing intact CSDs (ΔA2-A4). Furthermore, washing of the filters in higher salt (0.5 M NaCl) resulted in complete stripping of label from fragments lacking a functional CSD (not shown), again emphasizing that binding by the basic tail domains is relatively salt-sensitive. Further analysis focussed on binding of riboprobe to proteins and their fragments which had not been previously denatured.

The intact FRGY2 proteins were crosslinked to riboprobe, fragmented by acid treatment, treated with ribonuclease and then separated by SDS–PAGE. Autoradiography revealed which of the fragments were in contact with the radio-labelled RNA. As shown in Fig. 5C (track 2), the FA-cleaved fragments containing isolated basic domains were crosslinked to the riboprobe. The best competitive combinations were again heparin (track 5) and poly(C,U) (track 7) in the presence of Mg2+. Note that poly(A,G) is a strong competitor when the intact protein with a functional CSD is involved (track 8). The use of HA-cleaved fragments in binding prior to crosslinking and separation confirmed that they behave as basic tail domains, i.e. riboprobe binding is competed by poly(C,U) and heparin and not by poly(A,G) in the presence of low concentrations of Mg2+ (not shown). Further FA digestion of the HA-cleaved material at DP sites generates individual labelled basic domains (Fig. 5D, tracks 2 and 3), which correspond to those produced by FA cleavage of FRGY2–riboprobe complexes (track 1). Again, binding is competed by heparin (track 4) and poly(C,U) (track 5) in the presence of Mg2+.

To what extent sequence recognition by the different binding domains of FRGY2 proteins is Mg2+-dependent rather than cation-sensitive was checked by substituting spermidine for MgCl2 in binding reactions prior to crosslinking. The results are summarized in Table 1 and lead to the main conclusions that: (i) The CSD interacts preferentially with poly(A,G), to a slightly lesser degree with poly(G), (not shown), but not at all with poly(A), this activity being specifically inhibited by 5 mM MgCl2; (ii) a low concentration of cations, 1 mM of either MgCl2 or spermidine, is required for a specific interaction of tail domains with poly(C,U), this activity being lost in 0.5 M NaCl or competed by heparin. As shown previously (30), poly(C,U) competes for protein binding, on a mass:mass basis, almost equally with the more complex riboprobe sequence. Similarly, in conditions favouring poly(A,G) binding, the synthetic ribopolymer competes almost as well as the riboprobe (not shown). However, the optimal recognition sequence, and hence its frequency of occurrence in the different riboprobes, is still unknown for each of the binding activities.

DISCUSSION

The procedure outlined here for the purification of Y-box proteins from mRNP could in principle be applied elsewhere, for example to extract Y-box proteins from somatic tissues, where they may be much less abundant than in *Xenopus* oocytes. It is assumed that the treatment employed does not disrupt the binding specificities of the proteins and, in particular, that the cold-shock domain is either stable at 80°C or is at least able to renature correctly. Thermostability appears to be a general property of Y-box proteins; for instance, the human YB-1 Y-box–DNA complex is resistant to heating to at least 69°C (3).

The initial observation was that the combined action of heparin and Mg2+ completely blocks binding of FRGY2 proteins to RNA. That Mg2+ influences the activity of Y-box proteins is apparent in other studies, notably with YB-1 (9), where it is suggested that Mg2+ enhances binding to apurinic DNA, and with unr (33) where binding to both DNA and RNA targets reveals a sensitivity towards Mg2+ at about 1 mM. The unr protein consists of a 5-fold repeat of the CSD with no auxiliary tail domains. It may be inferred that a similar sensitivity to Mg2+ applies to the single CSD of Y-box proteins. The effect of Mg2+, at least on FRGY2 binding to RNA, appears to be 2-fold: to interfere with binding by the CSD and to favour binding by the tail domains over the range 1–5 mM. These two effects appear to be due to different types of molecular mechanism, because the first is specific to Mg2+, whereas the second can be obtained with the structurally distinct cation spermidine. The contribution of heparin to the blocking of RNA binding is explained by the interaction of this polyanion with the arginine clusters contained in the tail domains, thus competing efficiently with RNA for binding sites.

To date, the ability of FRGY2 tail domains to bind RNA has been discussed (34), but not considered in detail. In the present study we show that after RNA binding, UV irradiation and protein fragmentation, individual tail domains of FRGY2 have been crosslinked to riboprobe. Furthermore, crosslinking of riboprobe

---

**Table 1. Effect of cations on the binding specificities of FRGY2 domains**

<table>
<thead>
<tr>
<th>Cations added:</th>
<th>Cold-shock domain</th>
<th>Basic tail domains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>poly(A,G)</td>
<td>poly(C,U)</td>
</tr>
<tr>
<td>poly(A,G)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(C,U)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(C,U)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(A,G)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(C,U)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(A,G)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(C,U)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(A,G)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>poly(C,U)</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Efficiency of crosslinking to riboprobe was measured in the presence of poly(A,G), poly(C,U), both at 40 μg/ml and the concentrations of NaCl, MgCl2 or spermidine shown. Scores of + represent competition levels of >90% of controls (minus competing polymer); scores of – represent competition levels of <20% of controls; and scores of (+) represent intermediate values. Scores of 0 indicate that no binding to riboprobe could be obtained under the conditions shown.*
to HA-cleaved fragments confirms that the tail domains can bind to RNA independently of a functional CSD. Loss of function of the CSD after treatment of FRG2Y proteins with HA is inferred from the marked increase observed in sensitivity of RNA binding to heparin and NaCl. The NG cleavage site lies within the $\beta$-strand of the CSD and cleavage would result in complete loss of the $\beta$-strand from the main part of the protein and disruption of the $\beta$-strand itself. Since the $\beta_1-\beta_3$-strands of the bacterial cold-shock proteins have been identified as forming the structure which interacts with nucleic acids (15-18) and since deletion of part of the $\beta_2$-strand of the NSEP-1 Y-box protein results in loss of DNA-binding activity (5), it is reasonable to assume that FRG2Y would similarly lose the function of its CSD on cleavage.

The reported binding specificity of Y-box proteins to DNA is remarkably varied. In addition to interaction of Y-box proteins with Y-box promoter elements of vertebrates (3,22,35,36), HTLV, HIV and RSV (37,38) and bacteria (19,20), there appears to be an affinity for pyrimidine-rich elements in the promoters of c-myc (5), $\gamma$-globin (7) and the liver-specific gene apoLV/DLI (6), and now also for purine-rich elements in the LTR of RSV (10). One possibility is that the different specificities observed are due to differences between Y-box proteins in regions less conserved than the essentially invariant CSD. Although the tail domains generally consist of alternating basic and acidic regions, their sequence is not conserved between different proteins. A second possibility is that the Y-box proteins all recognize a similar range of sequences but that different targets are recognized by different protein domains. This aspect has been investigated in the RNA-binding studies reported here, and similarities in the mode of binding of Y-box proteins to DNA and to RNA can be considered. For example, most of the studies made on the binding of Y-box proteins to DNA emphasize the preferred interaction with single-stranded targets; likewise, we note that double-stranded RNA presents a poor binding template. Furthermore, two Y-box proteins involved in developmental regulation, FRG2Y (22) and MSY1 (12,13), have important roles in masking mRNA from translation, in addition to their gene regulatory activities. Indeed, Y-box proteins in general may have a role in mRNA packaging.

In our RNA-binding assays we describe a preference for the polypurines p(A,G) and p(G), but not p(A), a binding mediated by the CSD. In addition we describe a cation-dependent preference for polypyrimidines due to the tail domains. This second type of binding has all of the characteristics that we described earlier when Mg$^{2+}$ was an integral component of the binding reaction (30). By analogy, the various DNA binding activities so far reported may be due to the activity of the CSD, tail domains or a cooperative combination of both, depending on the reaction conditions used. In the context of binding to promoter elements, the Y-box proteins may be associating with DNA via the cooperation of both types of domain. The presence of an actual Y-box DNA element appears not to be essential for binding to certain promoters where, for example, an H-DNA structure is induced in regions of strong purine—pyrimidine strand asymmetry, exposing a pyrimidine-rich single strand from a triplex strand (30). It is interesting to note that the structure of H-DNA appears to be stabilized by Mg$^{2+}$ (7), which is a factor determined here for activating tail domain binding to RNA. Mg$^{2+}$ is also known to stabilize secondary structures in RNA (40), raising the possibility that its influence on FRG2Y binding is through modifying the structure of the RNA itself. An alternative possibility is that Mg$^{2+}$ is involved directly in the interaction between amino acids and nucleotide residues, as has been suggested for the binding of HeLa transcription factor USF to the E-box (41).

Although the exact nature of the molecular interaction between Y-box proteins and nucleic acids is unknown, we have shown that in vitro binding to RNA is highly sensitive to ionic environment, so much so that the activity of the CSD is destroyed by 5 mM MgCl$_2$ and that the activity of the basic tail domains in binding polypyrimidines is dependent upon the presence of low concentrations of multivalent cations (Mg$^{2+}$ and spermidine$^{3+}$). Since the preference in sequence interaction is different for the two binding reactions, the mixed polypurine, poly(A,G), being favoured by the CSD and the mixed polypyrimidine poly(C,U) being favoured by the tail domains, a wide range of binding specificities could be achieved. Such versatility might be advantageous in packaging a wide variety of mRNA sequences, while at the level of promoter binding, different elements such as the Y-box sequence and CT-rich strands could be recognized under different in vivo conditions.

REFERENCES