Binding of the Protein Kinase PKR to RNAs with Secondary Structure Defects: Role of the Tandem A−G Mismatch and Noncontiguous Helixes†

Philip C. Bevilacqua,*,‡ Cyril X. George,§ Charles E. Samuel,§ and Thomas R. Cech‡

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado 80309-0215, and Department of Molecular, Cellular and Developmental Biology, University of California, Santa Barbara, California 93106-9610

Received January 15, 1998; Revised Manuscript Received March 9, 1998

ABSTRACT: The human interferon-induced double-stranded RNA (dsRNA)-activated protein kinase (PKR) is an antiviral agent that is activated by long stretches of dsRNA. PKR can also be activated or repressed by a series of cellular and viral RNAs containing non-Watson−Crick motifs. PKR has a dsRNA-binding domain (dsRBD) that contains two tandem copies of the dsRNA-binding motif (dsRBM). In vitro selection experiments were carried out to search for RNAs capable of binding to a truncated version of PKR containing the dsRBD. RNA ligands were selected by binding to His_6-tagged proteins and chromatography on nickel(II) nitrilotriacetic acid agarose. A series of RNAs was selected that bind either similar to or tighter than a model dsRNA stem loop. Examination of these RNAs by a variety of methods, including sequence comparison, free-energy minimization, structure mapping, boundary experiments, site-directed mutagenesis, and footprinting, revealed protein-binding sites composed of noncontiguous helices. In addition, selected RNAs contained tandem A−G mismatches ($^{5'AG3'}_{3'GA5'}$), yet bound to the truncated protein with affinities similar to duplexes containing only Watson−Crick base pairs. The NMR structure of the tandem A−G mismatch in an RNA helix $(rGGCAGGCC)_2$ reveals a global A-form helix with minor perturbations at the mismatch [Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) *Biochemistry 36,* 4449−4460]. This supports the notion that dsRBM-containing proteins can bind to RNAs with secondary structure defects as long as the RNA has an overall A-form geometry. In addition, selected RNAs are able to activate or repress wild-type PKR autophosphorylation as well as its phosphorylation of protein synthesis initiation factor eIF-2, suggesting full-length PKR can bind to and be regulated by RNAs containing a tandem A−G mismatch.

PKR is an interferon-induced human protein kinase that can regulate gene expression via multiple pathways. These include inhibition of translation initiation by phosphorylation of initiation factor eIF2α (*1*, *2*) and modulation of cytokine signaling and transcription activation by the NF-κB and STAT1 factors (*3*, *4*). As a result of its varied biochemical actions, PKR has been implicated as an antiviral and antiproliferative agent. For example, PKR has been shown to be a regulator of human immunodeficiency virus type 1 (HIV-1) replication (*5*−*7*) and adenovirus replication (*8*, *9*), as well as an inducer of apoptosis (*10*−*13*).

PKR contains a double-stranded RNA (dsRNA)-binding domain (dsRBD) that consists of two tandem copies of the conserved dsRNA-binding motif (dsRBM) (*14*−*18*). The dsRBM is a 65−68 residue, compact RNA-binding motif that occurs in a large number of functionally diverse proteins from a variety of organisms (*19*, *20*). The NMR structure of the dsRBM has been solved and has a secondary structure repeat α-*β*-*β*-*β*-α, consisting of two α-helices packed on the same face of a three-stranded antiparallel *β*-sheet (*20*, *21*). PKR binds to dsRNA but not RNA-DNA hybrids or dsDNA. This discrimination arises because the protein makes only one energetically significant ion pair with the phosphate backbone, which is similar between the dsRNA and RNA-DNA hybrids, and instead relies on a network of 2′-OH functional groups on both strands of the helix and along the length of the binding site (*22*).

By virtue of its dsRBD, PKR can bind to and be activated by dsRNA of viral origin in a sequence-independent fashion (*21*, *23*−*28*). The fact that viruses have evolved a variety of mechanisms for down-regulating PKR function, including production of RNAs and proteins that interfere with the activation pathway (*2*, *9*, *29*−*32*), further illustrates the importance of PKR. RNAs that regulate PKR contain numerous unpaired or non-Watson−Crick interruptions of helices, including the 3′-untranslated region (UTR) from human α-tropomyosin (*33*) and RNAs of various viruses that replicate in human cells. These include human hepatitis *δ* agent RNA (*34*, *35*), adenovirus VA I RNA (*9*, *36*), Epstein−Barr Virus EBER-1 and −2 RNAs (*37*, *38*), and HIV TAR RNAs (*39*−*41*). PKR can also associate with ribosomes (*24*), perhaps binding to non-Watson−Crick motifs of rRNA (*42*). Observation of PKR regulation by RNAs with second-

† This work is supported by a fellowship to P.C.B. from the Jane Coffin Childs Memorial Fund for Medical Research and by a research grant to C.E.S. from the National Institute of Allergy and Infectious Diseases. T.R.C. is an investigator of the Howard Hughes Medical Institute and a Professor of the American Cancer Society.

* To whom correspondence should be addressed. Present address: Department of Chemistry, The Pennsylvania State University, University Park, PA 16802.

‡ University of Colorado−Boulder.

§ University of California−Santa Barbara.

S0006-2960(98)00113-5 CCC: $15.00
Published on Web 04/16/1998

ary structure defects instigated this investigation into the general role of RNA structure and sequence in binding to the dsRBM.

An RNA selection approach was utilized to search for RNAs from a random library that are capable of binding to a $(His)_6$-tagged version of the dsRBD from PKR. A pool of 10^{14} RNA sequences containing a randomized region of 50 nucleotides was prepared, from which RNAs capable of binding to the dsRBD were selected. Structural analysis of these RNAs suggests a general role for non-Watson−Crick motifs, as well as noncontiguous helices, in dsRBD binding.

MATERIALS AND METHODS

PKR Protein Construct. Selection experiments were performed with H_6Tp20, an N-terminal, $(His)_6$-tagged protein (*22*) that contained residues 1−184 of the wild-type (Wt) 551-residue PKR protein (*43*). H_6Tp20 protein was overexpressed in *Escherichia coli* and purified to greater than 90% homogeneity by affinity chromatography with nickel(II) nitrilotriacetic acid agarose (*22*).

Selection Procedure. Selection procedures followed the basic approaches described (*44*, *45*).

(1) Preparation of the Randomized dsDNA Template. A double-stranded DNA fragment was prepared by PCR using top- and bottom-strand primers, TS2 and BS2, respectively, to amplify a DNA template, Random2: TS2, GGGGGAATTCTAATACGACTCACTATAGGGAGAGCGGAAGCGTGCTGGGCC; BS2, GGGGGGATCCATCGACCTCTGGCTTAAG; Random2, GCGGAAGCGTGCTGGGCC-N(50)-CTTAAGCCAGAGGTCGAT; where each N position is an approximately equimolar mixture of the four nucleotides. The resultant dsDNA fragment had a 50-nucleotide random region, a T7 promoter sequence, and *Eco*RI and *Bam*HI cloning sites. The PCR reaction was carried out in a total volume of 3.3 mL using 165 pmol, or approximately 10^{14} different sequences of the Random2 template. PCR reactions were heated to 95 °C for 1 min prior to addition of 16.5 μL of 5 units/μL Taq polymerase (Perkin-Elmer). The PCR reaction was then performed for five cycles of a three-temperature amplification (1 min at 95 °C, 1 min at 54 °C, and 1 min at 72 °C), using high primer concentrations of 5 μM each. This led to a single band of correct mobility on an agarose gel. Eight and more rounds of PCR led to slower and faster mobility artifact bands and were thus avoided. These artifacts have been attributed, and least in part, to formation of heteroduplexes (*46*).

Reactions were examined on a 5% NuSieve agarose gel (FMC BioProducts), which indicated approximately 460 pmol of a double-stranded DNA product 129 bp in length. The PCR reaction was extracted with 1 vol of chloroform:isoamyl alcohol (24:1), ethanol precipitated, and washed in 70% ethanol. The resulting pellet was stored in TEN_{50} [10 mM Tris (pH 7.5), 1 mM EDTA, and 50 mM NaCl].

(2) Preparation of the Randomized RNA Library. A T7 transcription reaction was performed as described previously (*22*, *47*), with the following exceptions. Transcription was in the presence of [α-^{32}P]ATP in a total volume of 1 mL, and half of the dsDNA product from the PCR reaction was used as a template. After reaction for 2 h at 37 °C, 60 μL of 0.5 M EDTA was added to chelate the Mg^{2+}. The solution was ethanol precipitated, dissolved in 100 μL of TE [10 mM Tris (pH 7.5) and 1 mM EDTA], and passed over a Sephadex G-25 spin column. The eluate was digested with 50 units of DNase I (Pharmacia) for 20 min at 37 °C. The resulting solution was phenol extracted, ethanol precipitated, washed twice in 70% ethanol, and redissolved in 100 μL of HE [10 mM Hepes (pH 7.0), 0.1 mM EDTA]. When this purified RNA was run on a denaturing 6% polyacrylamide gel with radiolabeled single-stranded DNA markers, there was one major RNA product close to the expected length of 102 nucleotides. RNA concentration was determined spectrophotometrically. This pool of RNAs is referred to as pool 0 RNA. The pool of RNAs transcribed *after* round *n* of selection is referred to as pool *n* RNA.

(3) Selection of the H_6Tp20-Binding RNAs. Selection steps were at 22 °C. Before selection for binding to H_6Tp20 protein, the RNA pool was cleared with nickel(II) nitrilotriacetic acid agarose resin to remove potential resin-binding RNAs, as follows. Internally ^{32}P-labeled RNA (100 μL) was renatured by heating at 95 °C for 3 min in TE and incubating at 22 °C for 10 min. Selection buffer (SxnB) was added [1 × SxnB: 150 mM NaCl, 25 mM Hepes (pH 7.5) and 5 mM 2-mercaptoethanol], followed by 40 μL of a 50% slurry of nickel(II) nitrilotriacetic acid agarose resin (Qiagen) previously equilibrated in 1 × SxnB. This mixture was manually shaken for 10 min at 22 °C to ensure good contact between the resin and the solution. The mixture was centrifuged for 10 s at 14 000 rpm in an Eppendorf centrifuge, and the supernatant was reserved. The resin was washed once with 100 μL of 1 × SxnB, and the wash and supernatant were combined. When the cleared resin was washed three times with 500 μL of 1 × SxnB (compared to 10 × 500 μL washes of the protein-bound resin), only 0.044% or less of the RNA remained bound to the resin. This indicates that inadvertent enrichment of resin-binding RNAs did not occur.

After the preclear steps, herring sperm DNA, as per Bevilacqua and Cech (1996), and H_6Tp20 were added at concentrations of 0.1 mg/mL and 1 μM, respectively. Binding was allowed to occur in solution for 5 min prior to introduction of 100 μL of a 50% slurry of nickel(II) nitrilotriacetic acid agarose resin in 1 × SxnB. This step allowed RNA to bind to a free protein rather than an immobilized one. Five minutes is sufficient time to permit complete binding of H_6-Tp20 to model duplex RNA (*22*). Binding was in a higher salt than in a previous study (*22*), 150 mM versus 10 mM NaCl, to increase the stringency of binding. Although the salt dependence for binding of a 20 bp model dsRNA to H_6Tp20 is shallow, the increase in salt concentration used here still corresponds to an ∼10-fold decrease in binding affinity (*22*).

The RNA-protein-agarose mixture was manually shaken for 10 min at 22 °C, followed by centrifugation as above. The supernatant was removed, and the resin was washed 10 times with 500−600 μL of 1 × SxnB for each wash. The RNA was eluted with two or three 100 μL portions of 1 × elution buffer [ElB: 150 mM NaCl, 25 mM Hepes (pH 7.5), 5 mM 2-mercaptoethanol, and 250 mM imidazole]. Imidazole can directly compete with histidine for Ni^{2+}-binding sites. EDTA was not chosen as the elution agent since it has the potential for directly eluting RNAs with high Ni^{2+} affinity and leading to a false positive. The eluate fractions were pooled and phenol extracted. Glycogen carrier (20 μg from Boehringer Mannheim) was added, the RNA was

Table 1: Progress of in Vitro Evolution[a]

round	[RNA] (μM)	[protein] (μM)	%protein[b] bound
1	11	5	3.3
2	1	5	2.5
3	0.7	1	2.1
4	1	1	1.5
5	1.6	1	6.9
6	1.7	1	0.95
7	1.8	1	2.2
8	2.6	1	2.2
9	3	1	2.9
10	3.4	1	5.5
11	1.6	0.33	5.2
12	2.7	0.33	1.9

[a] The competitor was ssDNA for round 1, tRNA for all subsequent rounds. [b] Percent protein bound is equal to the percent RNA eluted times the fold-excess of RNA concentration over protein concentration. Liquid scintillation counting was used to determine the amounts of eluted and total RNA. Percent RNA eluted is equal to the ratio of the amount of the eluted RNA to the amount of the total RNA, $\times 100$.

ethanol precipitated and washed with 70% ethanol. The pellet was dried in a speed-vac, dissolved in 50 μL of TE, and passed over a Sephadex G-25 spin column. The RNA was ethanol precipitated, washed with 70% ethanol, dried in a speed-vac, and dissolved in 10 μL of TE. Fractions were collected at all steps and quantitated via liquid scintillation counting.

(4) RT-PCR To Convert the Selected RNA into dsDNA Template. Selected RNA (4 μL) was annealed with BS2, and cDNAs were obtained by reverse transcription for 30 min at 60 °C using AMV reverse transcriptase (Life Sciences). To this solution was added 50 μL of a master mix containing buffer and TS2. This mixture was heated at 95 °C for 3 min prior to addition of Taq polymerase. PCR was performed for 10 cycles, each involving the two-temperature amplification: 1 min at 95 °C and 1 min at 72 °C. The products of the RT-PCR reaction were examined on a 5% agarose gel. Two control reactions were carried out in parallel with each RT-PCR reaction. One control was without addition of selected RNA to test for DNA contamination of the stock solutions. The other control was with 1 μL of RNA added but without any reverse transcriptase to test for DNA contamination of the selected RNA.

(5) Later Rounds of Selection. Subsequent rounds of selection were the same as described for round 1 with the following exceptions. T7 transcription reactions were in 100 μL volumes and loaded directly on the Sephadex G-25 spin columns without any precipitation. The DNase step was not performed at this time, but rather after selection (see below). The RNA was refolded at 95 °C for only 1 min, and the competitor was 0.1 mg/mL ($\sim$6 μM) $tRNA^{Phe}$, not 0.1 mg/mL single-stranded (ss) DNA, to increase the stringency of binding (*22, 48*). tRNA can compete 2−8-fold better than ssDNA for H_6Tp20 binding to TAR and dsTAR, respectively (*22*). The concentration of H_6Tp20 was lowered in later rounds to increase the stringency of binding (see Table 1). After the selected RNA was ethanol precipitated with glycogen carrier, it was subjected to a DNase digestion. This digestion was in 40 μL at 37 °C for 2 h with 40 units of RQ1 RNase-free DNase (Promega). The DNase step was moved so that it followed the selection, thereby reducing the amount of nucleic acid present during the digestion. After the DNase step, the solution was phenol extracted. Glycogen (20 μg) was added, and the RNA was ethanol precipitated and washed with 70% ethanol. The pellet was dried in a speed-vac and dissolved in 10 μL of TE. PCR steps involved 15−23 rounds of the above two-temperature amplification.

After 12 rounds of selection, the RNA was converted into cDNA by reverse transcription and amplified by PCR. The PCR fragments were digested with *Eco*RI and *Bam*HI and cloned into the plasmid pUC19. Sequences of isolated clones were determined by dideoxy sequencing, using sequencing gels that contained 40% formamide and 7 M urea in order to resolve compressions.

Preparation of RNAs and DNAs. RNA was prepared from desired clones by runoff T7 transcriptions. The appropriate plasmid was digested with *Bam*HI and transcribed as described previously (*22*). The reaction was ethanol precipitated, run on a 6% denaturing polyacrylamide gel, and product bands visualized by UV shadowing. RNA was recovered by a crush and soak method (*22*).

All other RNA and DNA were prepared by solid-phase synthesis and deblocked as previously reported (*49, 50*). RNA sequences are found in the appropriate figure or figure caption.

Binding Assay. Binding affinity between H_6Tp20 and 5′-^{32}P-labeled RNA was determined by native-gel mobility-shift assays, as previously described (*22*). Fits are to eq 1,

$$\theta = \epsilon \frac{[\mathrm{PKR}]}{[\mathrm{PKR}] + K_\mathrm{d}} \qquad (1)$$

where θ is the fraction of RNA bound, ϵ is the observed maximum fraction bound, and K_d is the dissociation constant (*22*). As necessary, complexes 1 and 2 were summed to give total RNA bound and obtain a K_d for binding of one protein to free RNA. This appears justified because complex 2 does not begin to form until essentially all of the free RNA is shifted. This suggests that binding of the first and second proteins are not strongly coupled, and that conversion of complex 1 into complex 2 does not effect the observed K_d for complex 1. Some of the gel mobility-shift data was cooperative and not well-described by a simple hyperbola. The origin of this effect is not well understood. In these cases, the K_d reflects the concentration of protein required to give half-maximal RNA bound.

Structure Mapping. RNAs prepared by runoff T7 transcription were treated with calf intestinal phosphatase (CIP) to remove the 5′-triphosphate, reacted with polynucleotide kinase and [γ-^{32}P]ATP, repurified by gel electrophoresis, excised from the gel, eluted overnight in TEN_{250} at 4 °C, ethanol precipitated, and resuspended in TE. 5′-Labeled RNA in TE was renatured by heating at 95 °C for 1 min and incubating at 22 °C for 10 min. RNA was digested with the appropriate nuclease for 1 h at 22 °C, under native conditions in 1 × SxnB. Nuclease concentrations were chosen to give limited hydrolysis, and 1 x concentrations were 1×10^{-5} units/μL RNase T1, 2×10^{-3} units/μL RNase U2, 2×10^{-6} units/μL RNase A, 1×10^{-4} units/μL RNase T2, and 0.07 units/μL RNase V1. Samples were either loaded immediately on a denaturing polyacrylamide gel or frozen on dry ice, thawed on ice, and loaded on a gel. Samples treated with imidazole were as previously described (*51*). Briefly, a 10 μL aliquot of 5′-end labeled RNA was mixed with an equal volume of 4 M imidazole (pH 7.0) and

incubated at 22 °C for 16 h. A 10 μL portion of this mixture was heated at 90 °C for 10 min as a denaturing control. To each aliquot were added 10 μL of 0.3 M NaOAc (pH 5.0) and 200 μL of a 2% $NaClO_4$ solution in acetone. The RNA was pelleted by centrifugation, washed with 400 μL of acetone, air-dried, and dissolved in 10 μL of TE followed by 10 μL of a 2× formamide loading buffer.

Sequencing lanes for A and G were prepared by limited hydrolysis with RNases U2 and T1, respectively, under denaturing conditions. Control lane was same as the G-sequencing lane with RNase T1 omitted. All-nucleotide lanes were prepared by treatment with alkali (*52*).

Footprinting Experiments. RNAs were 5′-end labeled, purified, and renatured as in the structure-mapping experiments. An aliquot of renatured RNA was incubated for 5 min at 22 °C with 7.7 μM H_6Tp20, and then incubated with the appropriate nuclease for 30 min at 22 °C. Samples were placed on ice, mixed with an equal volume of a formamide/0.5% SDS loading buffer, and 1 μL was loaded on an 8% polyacrylamide gel.

Boundary Experiments. RNA was either 5′-end labeled as in the structure mapping experiments or 3′-end labeled with poly(A) polymerase and [α-^{32}P]cordycepin-5′-triphosphate as per manufacturer's instructions (United States Biochemical). Labeled RNAs were treated with alkali to generate a ladder of sequences. Hydrolyzed RNAs were mixed with 5 μM H_6Tp20 and loaded on a native gel, as previously described (*22*). Bound and unbound RNAs were visualized by autoradiography, excised from the gel, and eluted by a crush and soak method. RNAs were ethanol precipitated and run on an 8% denaturing polyacrylamide gel, along with T1 and alkali digests.

In Vitro PKR Activation Experiments. The effect of H_6-Tp20-selected RNAs on PKR kinase activity was determined by measurement of both PKR autophosphorylation and eIF-2 phosphorylation, as previously described (*53*, *54*). PKR-(Wt) enzyme was purified from ribosomal salt-wash fractions prepared from interferon (IFN)-treated human amnion U cells as described (*53*), except that the procedure was modified to include a MonoQ ion-exchange fractionation step (*54*). [γ-^{32}P]ATP-mediated PKR autophosphorylation catalyzed by PKR(Wt) was carried out in the absence of added activator or in the presence of RNA as indicated. For measurement of eIF-2α phosphorylation by PKR, 0.25 μg of purified eIF-2 was added to the standard reaction mixture. The ^{32}P-labeled products were analyzed by SDS−polyacrylamide gel electrophoresis and autoradiography. The procedures have been described in detail (*53*). Quantitation was carried out using a Bio-Rad Molecular Imager model GS525.

RESULTS

Use of Ni^{2+}-Affinity Chromatography to Select Protein-Binding RNAs. Several control experiments were performed to determine optimal conditions for selection experiments. Binding of complexes of H_6Tp20 with [5′-^{32}P]TAR or with dsTAR complex (*22*) to nickel(II) nitrilotriacetic acid agarose resin in batch (Materials and Methods) was compared to binding in a silica-based spin-column (Qiagen). [dsTAR is a double-stranded version of TAR with a 24 bp stem in which the three TAR bulges are deleted and G:U wobble pairs are converted to G:C base pairs (*22*).] In general, 10% more of the ^{32}P-labeled RNA was bound in batch compared to spin-column mode, as determined by liquid scintillation counting (data not shown). More importantly, in batch mode trials using only two 600 μL washes, only 0.1% of the RNA was bound and then eluted from the column without protein present, compared to 17−33% of the RNA with protein present. For the spin columns, 3−7% of the RNA was bound and then eluted from the column without protein present, compared to 10−23% of the RNA with protein present. Such a high background of RNA binding to the spin column in the absence of protein was unacceptable.

During each of the batch-mode selection steps, the protein dependence of binding was determined by comparing the counts per minute of the 300 μL of 250 mM imidazole-containing eluate to the counts per minute of the 10th and final 500 μL wash of the RNA-protein-agarose mixture. Between 91 and 98% more counts per minute eluted upon addition of the imidazole-containing buffer, indicating that the elution was specific to protein-bound RNAs. Also during each of the batch-mode selection steps, the efficiency of the elution was determined by comparing the counts per minute of the 300 μL of eluate to the counts per minute of a suspension of the used resin. This percent efficiency varied from 92 to 99% depending on the round of selection. In conclusion, nickel(II) nitrilotriacetic acid agarose resin appears to provide a generally useful approach for selection of RNAs that bind to a $(His)_6$-tagged protein. When used in batch mode, the resin does not significantly bind nucleic acids, while specific and efficient elution of protein and its bound RNAs is achievable.

Selection for RNAs That Bind the dsRBD from PKR. A library of $\sim 10^{14}$ different sequences was generated by overlap extension of top- and bottom-strand primers with a DNA template that had a stretch of 50 randomized positions, followed by transcription with T7 RNA polymerase. The renatured RNA pool was cleared with nickel(II) nitrilotriacetic acid agarose resin and recovered. The RNA was bound to protein in the presence of a competitor and washed exhaustively. RNA was eluted from the resin by several incubations with an imidazole-containing buffer.

Multiple rounds of selection were performed. The stringency of the binding step was increased at several points during the selection (Table 1). After 12 rounds of selection, the RNA was 5′-end labeled and its affinity for H_6Tp20 was quantitated by a native-gel mobility-shift assay (Figure 1). This heterogeneous collection of RNAs bound H_6Tp20 with an apparent K_d of 0.1 μM, 8- and 64-fold more tightly than dsTAR and TAR, respectively (Figure 1). Controls showed that pool 0 RNAs did not bind H_6Tp20 over the same range of protein concentration. At this point, selection was stopped and sequences were determined.

The RNA was converted to cDNA by reverse transcription, amplified by PCR, cloned, and sequenced. Of the 44 sequences determined, 11 were different, and five of the 11 had a 17-nucleotide consensus, "cugcaucgacaryu-ggy" where r is one of the purine bases A or G, and y is one of the pyrimidine bases C or U. These five sequences were assigned to family 1 (Figure 2). Many of the selected RNAs had deletions in the randomized region that presumably arose during the RT-PCR steps (*55*). We chose clones 2 and 11, the most frequent members of family 1, to characterize further. RNA derived from other clones, including clone 3,

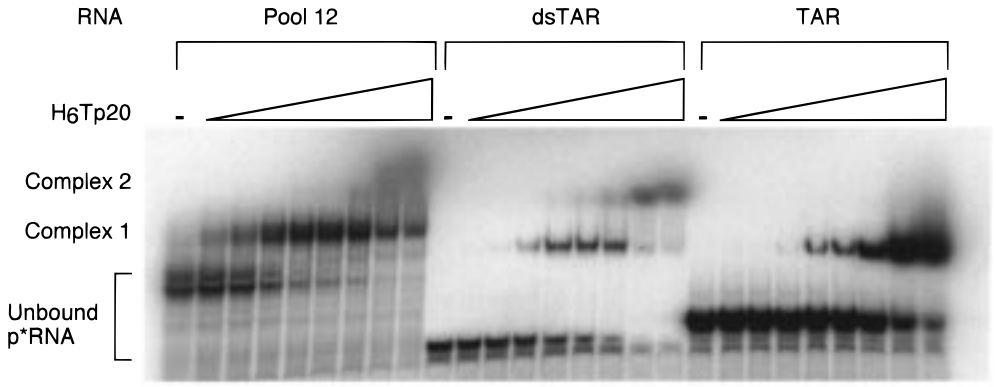

FIGURE 1: Native gel mobility shift for H_6Tp20 binding to RNA. Native-gel mobility-shift experiment for H_6Tp20 binding to trace amounts of 5′-^{32}P-labeled RNA. Pool 12 is the mixture of RNAs after 12 rounds of selection; TAR and dsTAR are partially mismatched and fully double-stranded RNA substrates, respectively, for H_6Tp20 binding as previously reported (*22*). Experiments were in the presence of 0.1 mg/mL $tRNA^{Phe}$. Native 10% polyacrylamide gel is shown. Protein binding to dsTAR resulted in slight formation of a second complex termed complex 2. Concentrations of H_6Tp20 used were 0, 0.05, 0.1, 0.25, 0.5, 1, 2, 5, and 10 μM.

```
                CONSENSUS      cugcaucgacARyu-ggy
Clone                                                          Number   Kd (μM)
Family 1
c2:   caugagguc-uau---cacugcaucgacAGcuaggcggacaaug               7       2.3
c20:  caugaggucguaugcgaucugcaucgacAGcuaggcggagcuaucg             1       4.1
c11:  ucggaggugg-uucagcucugcaucgacAGuu-ggc                       11      5.3
c14:     guaaggg-uuca-cucugcaucgacAAcu-gga                       3       6.2
c4:       uguuggcuugg-cucugcaucgacAGgu-gguu                      2       6.6

c3:   aggauggaucgagcgcuggagucuggagaaugucac                       6       0.074
c6:   ugcgauaaggugguagcgacagugauagaagc                           4       2.8
c5:   ugcacggaaaugcagaaggcgugguccgugggaggcgacaagcaucaaug         4       0.11
c21:  cacacgguggg-uguguauuacagcugacuguuggcaagcacugugaggac        3       N.D
c32:  uauacgguggggaugugcccaguggacacaugcaccguguuaguuuugca        2       N.D
c26:    cuuaacgugaaggcgguaggggugacggacugagcgagcgccaugcauug       1       N.D
```

FIGURE 2: Sequence alignment of selected RNAs. Shown are sequences from the variable region only. Gaps in the alignment are represented by a dash. Numbers are assigned to clones based upon the order in which the first occurrence was sequenced. The number of identical clones found in the 44 sequences is shown on the right. The consensus stretch of 17 nucleotides is shown at the top of the figure and is boxed. The first five clones have an obvious match to the consensus and are assigned to family 1. The other clones do not have an obvious match. Another region of similarity in these clones is shown in boldface and lowercase. Y = C or U, R = A or G, and the selected region's contribution to the tandem mismatch is shown in bold capital letters. K_ds are in units of micromolar and were determined by a native-gel mobility-shift method.

bound H_6Tp20 more tightly than those of family 1 (Figure 2). Clone 3 RNA, however, did not readily afford an H_6Tp20 footprint and was not easily classified by comparative sequence analysis; consequently, its structure was not pursued in this study.

RNA derived from clones 2 and 11 bound to H_6Tp20 both on the column and by gel shifts. Dissociation constants for clones 2 and 11 RNA binding to H_6Tp20 by native-gel mobility-shift experiments are 2.3 and 5.3 μM, respectively (Figure 2). These values are similar to that for dsTAR of 0.9 μM (Figure 1).

Effect of Selected RNAs on PKR Kinase Activity. The PKR kinase purified from human cells is highly dependent upon dsRNA for activity as measured both by autophosphorylation of PKR and by the phosphorylation of protein synthesis initiation factor eIF-2 by activated PKR (*53*, *54*). The ability of RNA from clones 2, 3, 11, and 14 to activate PKR was examined (Figure 3). Among these RNAs, clone 2 RNA activated PKR most efficiently. Although clone 3 RNA bound PKR more tightly than clone 2 RNA, clone 3 RNA was a much poorer kinase activator than clone 2 RNA. Optimal PKR activation with clone 2 was obtained between 0.1 μM and 1.0 μM RNA; this activation by clone 2 RNA approached the maximal activation of PKR obtained with synthetic poly(rI)-poly(rC) dsRNA (Figure 3A). Family 1 RNAs differed in their ability to activate PKR in the order clone 2 > clone 14 > clone 11. Relative to the activation of PKR obtained in the presence of clone 2 RNA, 54% activation was obtained with clone 14 RNA, 17% with clone 11 RNA, and 14% with clone 3 RNA (all RNAs tested at 1.0 μM).

To test whether selected RNAs could antagonize the activation of PKR kinase by dsRNA, purified PKR was incubated in the presence of selected RNA and then the reaction mixture was supplemented with 0.3 μg/mL (=2.4 nM) poly(rI)-poly(rC). As shown by the Figure 3B autoradiogram, clone 3 RNA was an effective antagonist of PKR activation by poly(rI)-poly(rC). About 90% reduction in PKR activation was obtained with 10 μM. Neither clone 2 nor clone 14 RNA significantly antagonized PKR activation and autophosphorylation relative to clone 3 RNA, whereas clone 11 displayed an intermediate antagonist activity.

Sequence Alignment and Free-Energy Minimization of Selected RNAs. The 17-nucleotide consensus sequence common to family 1 clones (Figure 2) can form 11−12 Watson−Crick or G−U wobble pairs with the fixed sequence 3′-primer-binding site. The proposed base-paired region is interrupted by an internal tandem A−G mismatch (with clone 14 forming an internal-5′AA3′/5′AG3′ mismatch). Each A−G pair in a tandem A−G mismatch can form an imino two hydrogen-bonded conformation, with A-NH6 to G−O6 and A−N1 to G-NH1 (imino) (*56*), which when added to

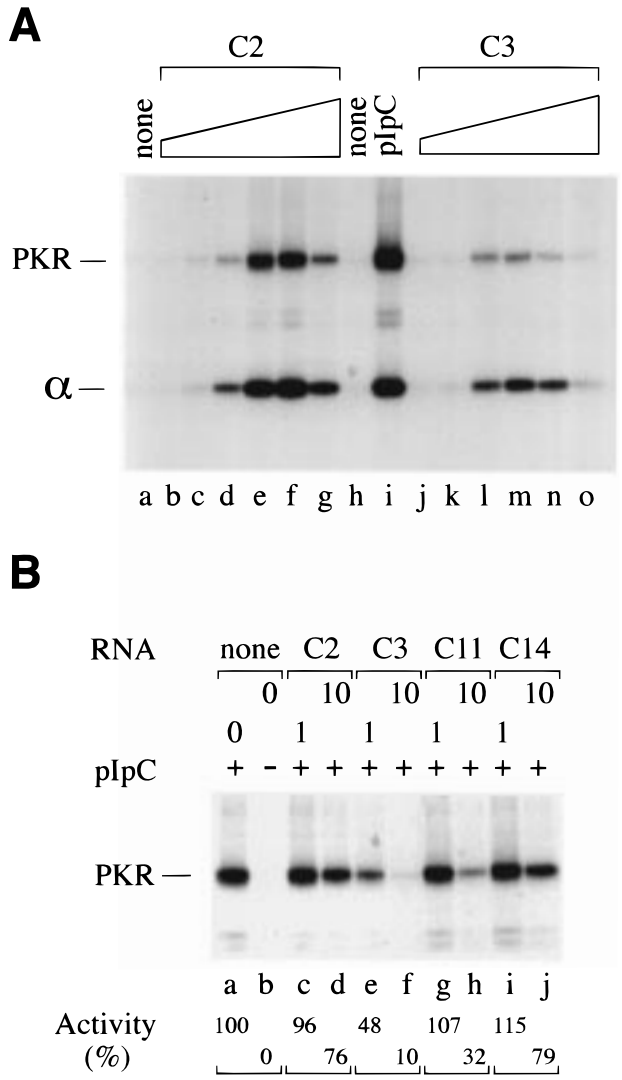

FIGURE 3: Effect of Selected RNAs on PKR Kinase Activity. (A) In vitro activation of PKR by selected RNAs. Shown is an autoradiogram of a 10% SDS−polyacrylamide gel. RNA samples, either clone 2 (lanes b−g) or clone 3 (lanes j−o), were examined for their ability to activate purified PKR kinase. RNA was added as indicated to the standard reaction mixture which contained [γ-^{32}P]-ATP and eIF-2. Concentrations of clone 2 and 11 RNAs used were: lanes b and j, 0.0001 μM; lanes c and k, 0.001 μM; lanes d and l, 0.01 μM; lanes e and m, 0.1 μM; lanes f and n, 1.0 μM; lanes g and o, 10 μM RNA. Reactions carried out in the absence of RNA (lanes a and h) or in the presence of the optimal concentration of poly(rI)-poly(rC) (lane i) are shown as controls. PKR denotes the position of ^{32}P-labeled PKR, and α denotes the position of ^{32}P-labeled α-subunit of eIF-2. (B) Antagonism of PKR activation by selected RNAs. Shown is an autoradiogram of a 10% SDS−polyacrylamide gel. PKR was incubated with RNA samples of clone 2 (lanes c and d), clone 3 (lanes e and f), clone 11 (lanes g and h), and clone 14 (lanes i and j) at either 1.0 or 10 μM prior to the addition of 0.3 μg/mL (=2.4 nM) poly(rI)-poly(rC). Reactions carried out in the absence of added RNA (lane b) or in the presence of the optimal concentration of poly(rI)-poly(rC) (lane a) but without added selected RNA, are shown as controls. PKR denotes the position of ^{32}P-labeled PKR. The autophosphorylation activity of PKR kinase was quantified and is expressed as a percent of the activity observed in the presence of an optimal concentration of poly(rI)-poly(rC).

the flanking helical regions would give a dsRNA-like helix of length 13 bp for RNA clones 4 and 14 and 14 bp for RNA clones 2, 11, and 20. This is close to the minimal length for binding of 16 bp identified in a previous study (*22*). These data suggest the model that the 3′-stem-loop of the family 1 RNAs forms the structure shown in Figure 4, containing the tandem A−G mismatch.

Secondary structures of various RNA clones were predicted by free-energy minimization using *Mfold* v2.3 (*57, 58*). Clone 11 RNA was predicted to form the structure in Figure 4 containing two major stem loops, P2/L2 and P3/L3. We show below that structure mapping, boundary experiments, site-directed mutagenesis, and footprinting experiments are consistent with this predicted secondary structure.

Clone 2 RNA was also predicted to form a structure with two major stem loops, P2/L2 and P3/L3. However, the 3′-stem-loop predicted to have the lowest free energy structure (−29.4 kcal/mol) has a structure different from that shown in Figure 4. The computer-predicted structure has an asymmetric loop with a top strand A opposing a bottom strand 5′CAGAG3′. Use of a window size of 0 in *Mfold* (*57*) identifies a suboptimal secondary structure only 0.7 kcal/mol less stable (−28.7 kcal/mol) with the tandem A−G mismatch structure shown in Figure 4. The lowest free-energy structure positions a nonconserved A in a stem base pair, while the suboptimal structure positions this A in a bulge as shown in Figure 4. It is most consistent to have a nonconserved residue not making an important interaction, further supporting the tandem A−G-containing structure of Figure 4.

Structure Mapping of Selected RNAs. To test whether the selected RNAs fold into the secondary structures suggested by sequence alignment and free-energy minimization, the RNAs were subjected to structure mapping. Cleavage of RNA clones 2 and 11 under native conditions by RNases U2, T1, A, T2, and V1, and by imidazole, is shown in Figure 5 and summarized in Figure 4. RNases U2, T1, A, and T2 are single-strand-specific nucleases that leave 3′-phosphate or 2′,3′-cyclic phosphate termini (*59*). RNase V1 is a double-strand specific nuclease that cleaves RNA without base specificity to leave a 5′-phosphate terminus (*59*). RNase V1 can also cleave single-stranded regions that exist in a stacked arrangement (*59, 60*). High concentrations of imidazole (pH 7.0) preferentially cleave RNA in single-stranded regions, leaving a 3′-phosphate terminus (*51, 61*). Positions that were cleaved by both single-stranded-specific RNases and by RNase V1 are not assigned a specific cleavage arrow in Figure 4. These ambiguous cleavages may be caused by single-stranded residues that are stacked, or by two competing RNA conformations that are in equilibrium.

The RNase and imidazole cleavage patterns were largely consistent with the predicted secondary structures (Figure 4). RNase V1 cleaved the selected RNAs within predicted double-stranded regions or at nucleotides one to two positions 3′ of double-stranded regions. A 3′-dangling end is able to stack on its proximal helix (*62*), consistent with its cleavage by RNase V1. Interestingly, RNase V1 cleaved clone 11 RNA lightly after position A52 and G53 of the tandem A−G mismatch. Single-strand-specific RNases and imidazole cleaved the selected RNAs within predicted single-stranded regions or at nucleotides involved in base pairs at the end of a stem. Such base pairs may be partially frayed, leaving them susceptible to attack.

As mentioned, the structure for clone 2 RNA shown in Figure 4 is a suboptimal structure according to calculation. The lowest free-energy structure predicted for clone 2 RNA

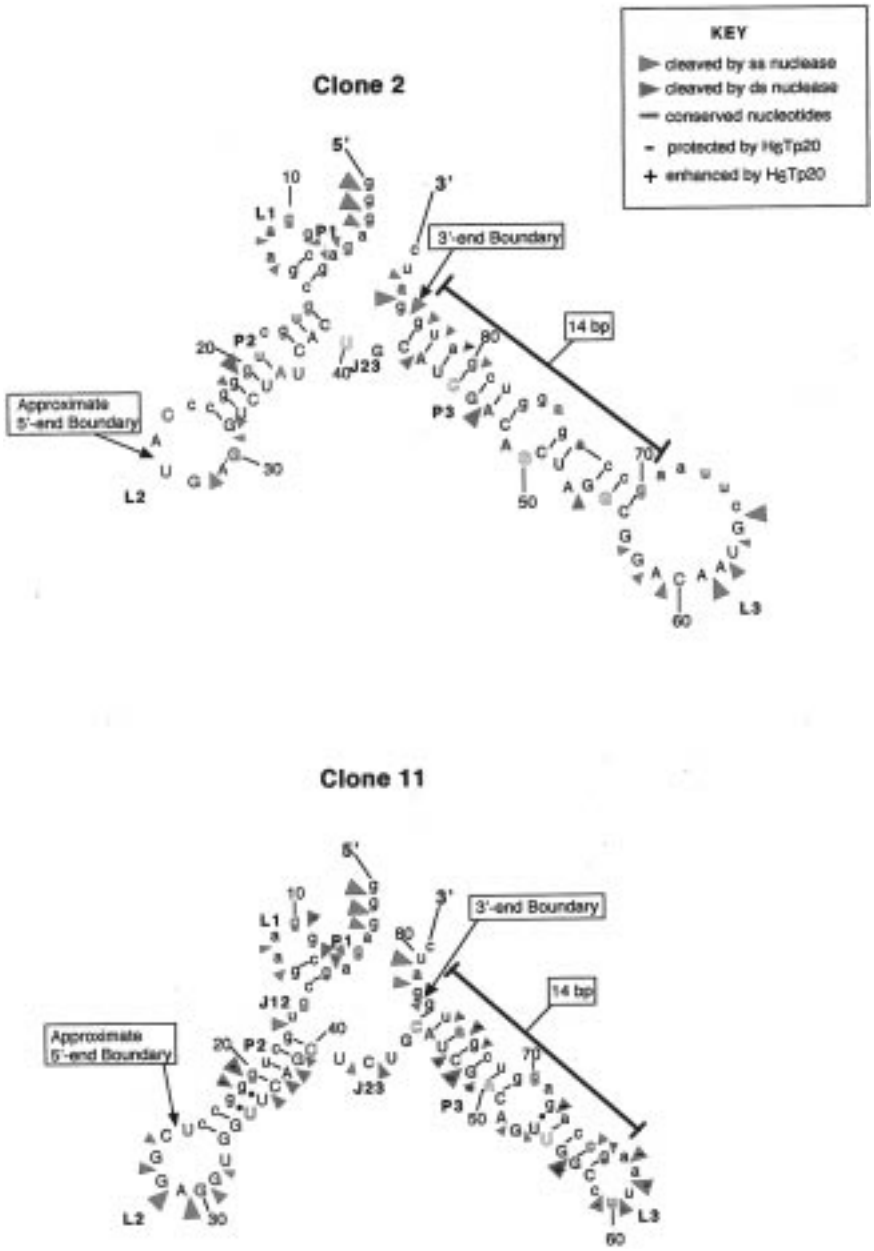

FIGURE 4: Secondary structural models for selected RNAs. Positions of cleavage by single-stranded probes (RNase U2, RNase T1, RNase A, RNase T2, and imidazole) are shown using red arrowheads with the size of the arrowhead proportional to the intensity of the cleavage band. Positions of cleavage by the double-stranded probe RNase V1 are shown using green arrowheads with the size of the arrowhead proportional to the intensity of the cleavage band. Ambiguous sites are not marked with any arrowhead. Nucleotides are numbered every 10 residues with each fifth residue outlined. Positions of the 5′- and 3′-boundaries are shown, as are positions of H_6Tp20 protection and enhancement of nuclease cleavage. Conserved nucleotides are bold and colored purple, and the length of the 3′-stem is indicated. Length of the tandem-mismatch containing stem is labeled. Number of base pairs includes Watson−Crick base pairs denoted with a dash, G−U wobble pairs denoted with a dot, and the tandem mismatch. Fixed-sequence primer-binding sites are in lower case, and nucleotides from the random region are in upper case. Positions of nuclease cleavage protection and enhancement by H_6Tp20 are indicated with minus (−) and plus (+) signs, respectively. Paired, joining, and loop regions are denoted with a P, J, and L, respectively, and numbered in order of occurrence from the 5′- to 3′-direction.

had two inconsistencies with the cleavage pattern of Figure 4: positions 63 and 64, predicted to be double stranded (data not shown), were in fact cleaved by RNases T2 and T1, respectively. Thus, the nuclease-mapping data are consistent with the tandem A−G containing secondary structures shown in Figure 4, rather than the calculated optimal structure.

Boundary Experiments on Selected RNAs. Clone 2 and clone 11 RNAs were tested to define the minimal set of nucleotides required for protein binding. RNAs were either 5′- or 3′-end labeled, hydrolyzed to give a ladder of sequences, and mixed with protein. Bound and unbound RNAs were separated by native gel electrophoresis, concentrated, and run on a sequencing gel. 5′-End-labeled sequences show a well-defined cutoff between bound and unbound sequences with only four residues dispensable at the 3′-end (Figures 4 and 6). Identical experiments were carried out on clone 2 RNA (data not shown). This cutoff retains the 3′-stem-loop and confirms its importance in H_6-Tp20 binding.

3′-End-labeled sequences show a poorly defined cutoff between bound and unbound sequences, with bound sequences gradually decreasing in intensity with a decrease in

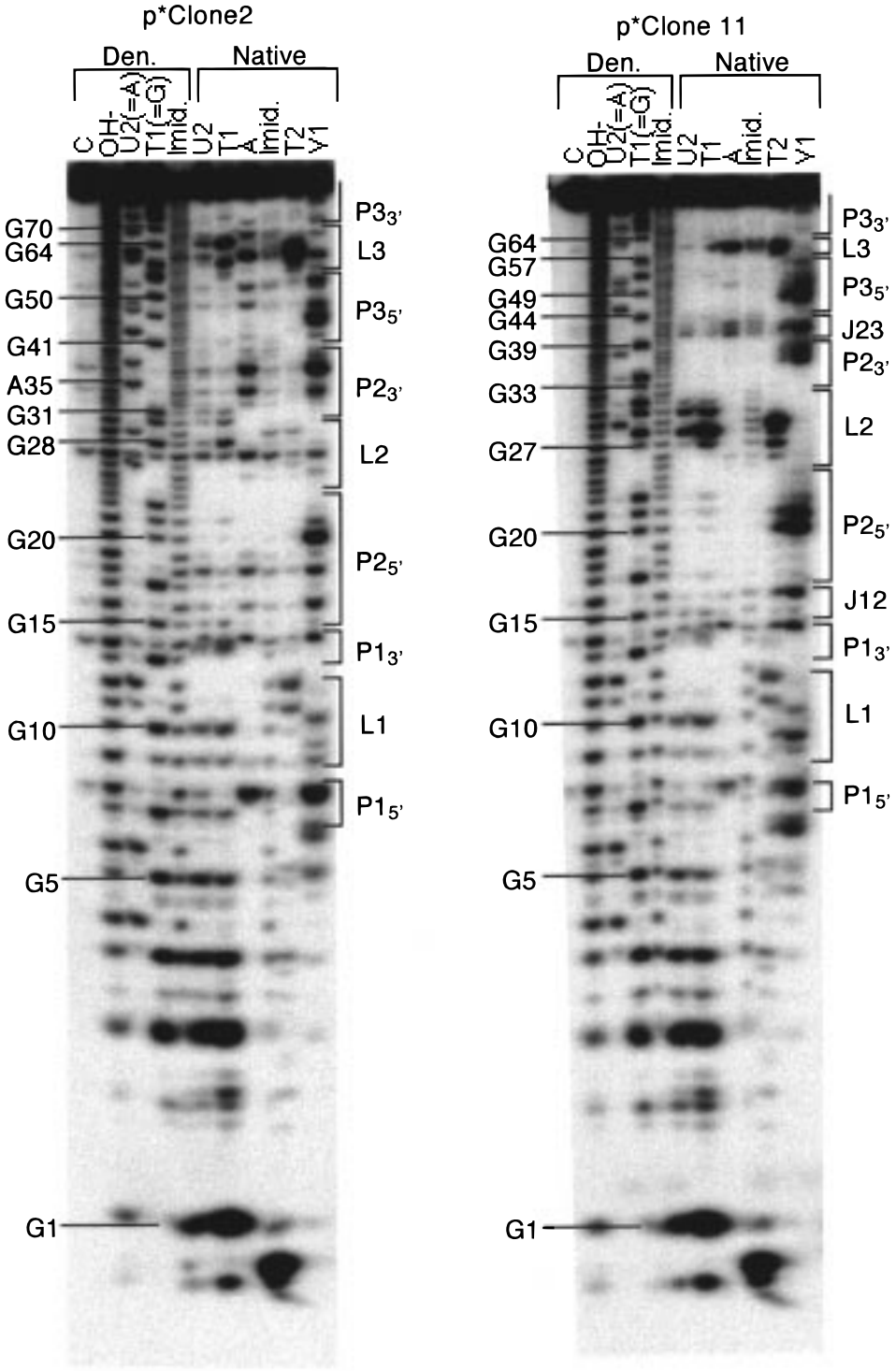

FIGURE 5: Structure mapping of selected RNAs. Denaturing 16% polyacrylamide gel is shown. RNA samples derived from clones 2 and 11 were 5′-^{32}P-labeled. Left-hand five lanes are under denaturing conditions (denoted "Den."), and right-hand six lanes are under native conditions (denoted "Native"). Labels for lanes are as follows: C is a control (no nuclease) sample, OH− is a limited alkaline digest; U2, T1, A, T2, and V1 are limited digests under native conditions with RNases U2, T1, A, T2, and V1, respectively; Imid. is a limited imidazole digest under native conditions. Samples were also run on a denaturing 8% polyacrylamide gel to help separate and identify the 3′-most cleavages (data not shown). Paired, joining, and loop regions are denoted with a P, J, or L, respectively, and the 5′- or 3′-strand of a region is denoted with a subscript.

RNA length and unbound sequences gradually increasing in intensity with a decrease in RNA length (Figure 6). As such, it is not possible to define precisely a 5′-end for RNA clone 2 or 11 that is absolutely required for binding. It is clear from the 3′-end labeled RNA boundary experiment, however, that the 14 bp 3′-stem-loop with the tandem A−G mismatch is not capable by itself of binding to H_6Tp20. The shortest sequences that are approximately 50% bound begin with residues 27 and 26 for RNA clones 2 and 11, respectively (Figure 6). RNAs derived from clones 2 and 11 that begin with residues 27 and 26, respectively, are capable of forming minihelices at their 5′-ends of 3 and 5 bp in length,

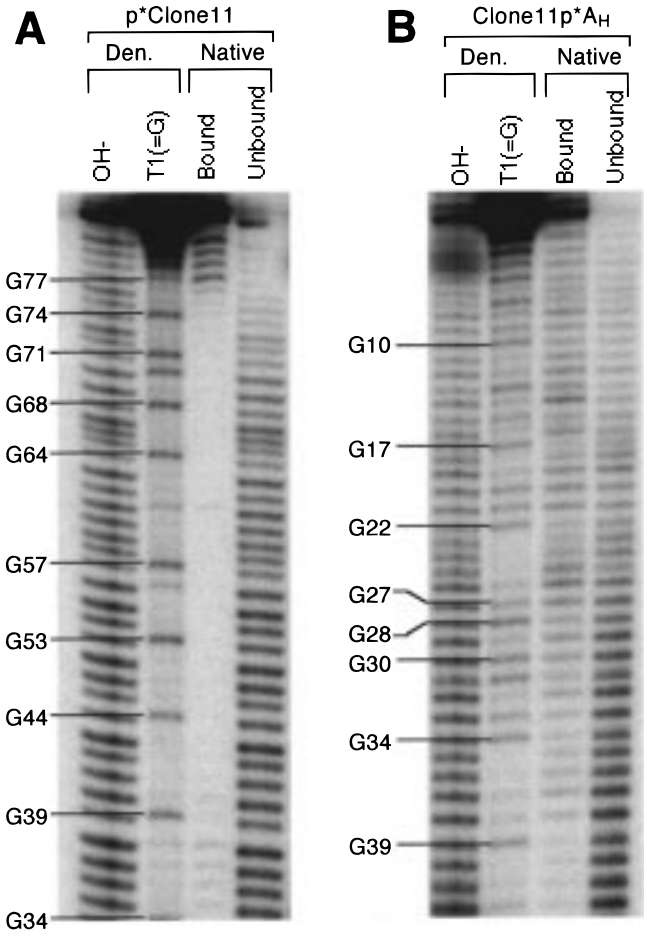

FIGURE 6: Boundary experiment for clone 11 RNA. Denaturing 8% polyacrylamide gel is shown. RNA was either (A) 5′-^{32}P-labeled, denoted "p*Clone11", or (B) 3′-^{32}P-cordycepin-labeled, denoted "Clone11p*A$_H$". Left-hand two lanes contain samples treated under denaturing conditions (denoted "Den."), and right-hand two lanes contain samples treated under native conditions (denoted "Native"). Labels for lanes are as follows: OH− is a limited alkaline digest; T1 is a limited digest with RNase T1; "Bound" and "Unbound" are RNA fragments from the limited alkaline digest that either bound or did not bind, respectively, to 5 μM H_6Tp20. Samples were also run a shorter distance on the 8% polyacrylamide gel to examine shorter fragments (data not shown).

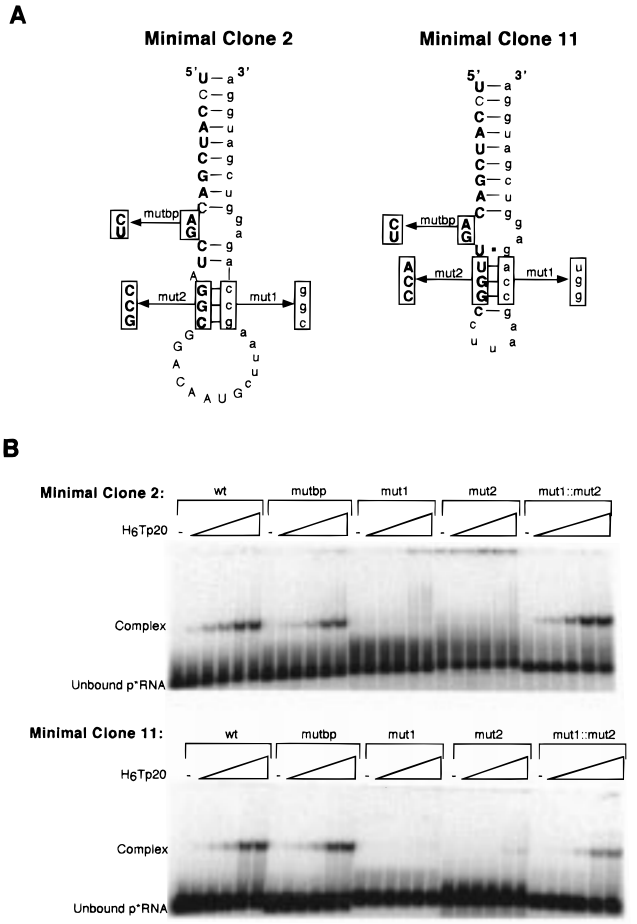

FIGURE 7: Structure and function of minimal-length RNAs. (A) Secondary structures for minimal RNAs derived from clones 2 and 11 consistent with structure mapping experiments, site-directed mutagenesis experiments, and free-energy minimization using *Mfold* v2.3 (*57*). Conserved nucleotides are bold. Positions of site-directed mutations in mutbp, mut1 and mut2 are indicated; mut1::mut2 is the double mutant with changes simultaneous for mutant 1 and mutant 2. (B) Native-gel mobility-shift experiments for H_6Tp20 binding to trace amounts of 5′-^{32}P-labeled RNA. Protein binding to RNA in the presence of 0.1 mg/mL tRNA^{Phe} resulted in formation of a single complex. Concentrations of H_6Tp20 used were 0, 0.5, 0.1, 0.2, 5, and 10 μM.

respectively, as evaluated by free energy minimization using *Mfold* v2.3 (*57*). These minihelices are predicted to fold in the presence of the 14 bp 3′-stem-loop structures shown in Figure 4. Since the 14 bp 3′-stem-loop itself cannot bind protein, and the minimal-length stem that can bind H_6Tp20 is known to be 16 bp (*22*), it is plausible that the 3−5 bp 5′-minihelices augment the A−G containing 3′-stem-loop stretch to form helical segments of sufficient length for protein binding. This is consistent with H_6Tp20 footprints present in both the 3′-stem-loop and the 5′-region of full-length clone 11 RNA (Figures 4 and 8, see below).

Minimal-Length RNAs: Structure and Site-Directed Mutagenesis. To test the secondary structure of the 3′-stem-loop and the importance of the tandem A−G mismatch in H_6Tp20 binding, a series of minimal RNAs and site-directed mutants was prepared. The minimal RNAs consist of the 3′-stem-loop with two additional base pairs at the end of the stem to give a 16 bp substrate (Figure 7A). Four mutant RNAs were also constructed: the base-pair mutant converts the tandem A−G mismatch of the model to two Watson−Crick base pairs, mutants 1 and 2 disrupt the stem proximal to the loop, and the double mutant combines mutants 1 and 2 to restore the proximal stem in the model (Figure 7A).

The wild-type, base-pair, and double-mutant RNAs were all capable of binding to H_6Tp20, while mutants 1 and 2 (in which the stem proximal to the loop was disrupted) were not capable of binding H_6Tp20 (Figure 7B). These results provide evidence that the secondary structures shown in Figure 7A are correct, that the tandem A−G mismatch is roughly equivalent to a Watson−Crick base pair in its ability to bind protein, and that the stem proximal to the loop is necessary for binding H_6Tp20. In addition, the sequence of the clone 2 double mutant, which is capable of forming the tandem A−G mismatch structure (Figure 7A), is not capable of forming the alternative secondary structure predicted by *Mfold* (*57*) to have an asymmetric loop. Since the double mutant retains H_6Tp20 binding capability, the secondary structure for clone 2 RNA shown in Figure 7A is capable of binding the protein.

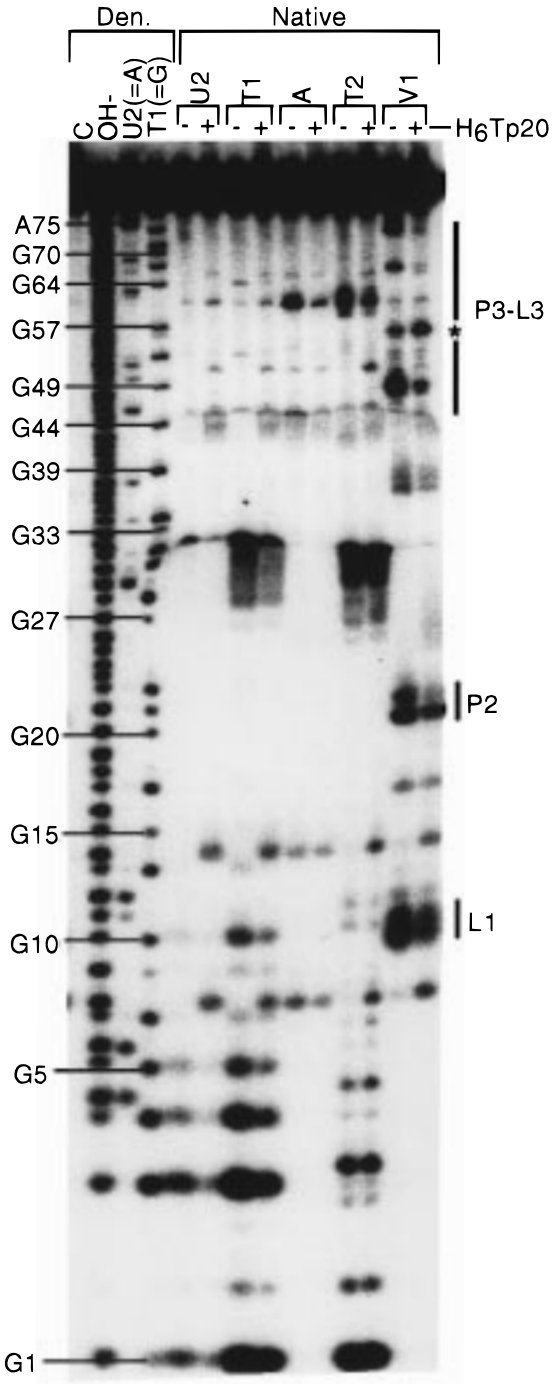

FIGURE 8: Footprinting experiment for clone 11 RNA. Denaturing 8% polyacrylamide gel is shown. RNA was 5′-^{32}P-labeled. Lanes are labeled as in Figure 5. In indicated lanes, H_6Tp20 was added at 7.7 μM (enough to give complete mobility shift of the complex prior to nuclease digestion). Nucleotides are labeled along the left-hand side of the gel. To the right of the gel, regions of H_6Tp20-dependent protection from nuclease cleavage are denoted with a bold line, and a site of H_6Tp20-dependent enhancement of cleavage is denoted with an asterisk (*).

Interaction of Selected RNAs with the dsRBD from PKR. To examine the binding sites for H_6Tp20 on the selected RNAs, footprinting experiments were performed. Treatment of the clone 2 and 11 RNAs with RNases U2, T1, A, T2, and V1 in the presence and absence of saturating H_6Tp20 is shown in Figure 8 and summarized in Figure 4. Positions that were cleaved by both single-stranded-specific RNases and by RNase V1 are not assigned a protection or enhancement. H_6Tp20 binding to clone 11 RNA gave a series of well-defined protections from RNase V1 in double-stranded sections, as well as one enhancement of RNase V1 cleavage at G56. In addition, protection from single-stranded nucleases occurred in L3.

H_6Tp20 binding to clone 2 RNA yielded only one well-defined protection, occurring with RNase V1 at position 80 in the 3′-stem. In sum, footprints map largely to the 3′-stem-loop of the structures shown in Figure 4, emphasizing the importance of this stem and its secondary structure.

Minimal-Length RNAs: Effect on PKR Kinase Activity. The relationship between RNA-binding activity of the minimal-length RNAs to H_6Tp20 and their ability to activate PKR kinase activity was examined. The clone 2 and 11 series of minimal RNAs and their site-directed mutants were tested at varying concentrations for their ability to activate PKR as measured by the phosphorylation of eIF-2α. As shown in Figure 9, the minimal-length clone 2 RNA activated PKR, but much less efficiently than the parental clone 2 RNA and only at relatively high concentrations. Minimal-length clone 2 mutant RNAs, like the minimal-length clone 2 RNA, were also very poor activators of PKR kinase activity (data not shown).

The ability of the minimal-length RNAs to antagonize the activation of PKR kinase by dsRNA was also examined. Purified PKR was incubated in the presence of minimal RNA, either wild-type or mutant, and then the reaction mixture was supplemented with poly(rI)-poly(rC). A series of clone 2 and 11 minimal RNAs and site-directed mutants was examined (Figure 10). As shown by Figure 10, minimal-length clone 2 base-pair mutant was a very effective antagonist of PKR activation. Mutant 1 minimal-length clone 2 RNA was a poor antagonist, whereas wt, mutant 2 and the double-mutant minimal-length clone 2 RNAs showed an intermediate ability to antagonize activation. By contrast, the minimal-length clone 11 series of RNAs, either wild-type or mutant, did not significantly antagonize PKR activation (data not shown).

DISCUSSION

The dsRBM is a conserved set of amino acids that constitutes a compact folding domain capable of binding dsRNA (*19*−*21*, *23*). In addition, dsRBM-containing proteins bind to RNAs containing single-stranded or non-Watson−Crick secondary structures (*9*, *28*, *33*−*41*, *63*). We have performed in vitro selection experiments to search for RNA sequences and structures that bind to the dsRBD from PKR. The selected RNAs might have been an uninteresting collection of double-stranded sequences. Instead the selected RNAs contain several non-Watson−Crick features, including a tandem A−G mismatch in the context of a dsRNA background, and a protein-binding site composed of non-contiguous helices. Thus, the results reveal new information about the binding specificity of the dsRBD from PKR.

Selected RNAs Bind H_6Tp20 and Activate PKR. After 12 rounds of selection, the pool 12 RNAs bound to H_6Tp20 approximately 8-fold more tightly than dsTAR, a model stem-loop containing 24 contiguous Watson−Crick base pairs (Figure 1). In addition, H_6Tp20 binding to selected pool 12 RNA clones was revealed by H_6Tp20-dependent protections from RNase cleavage (Figure 8).

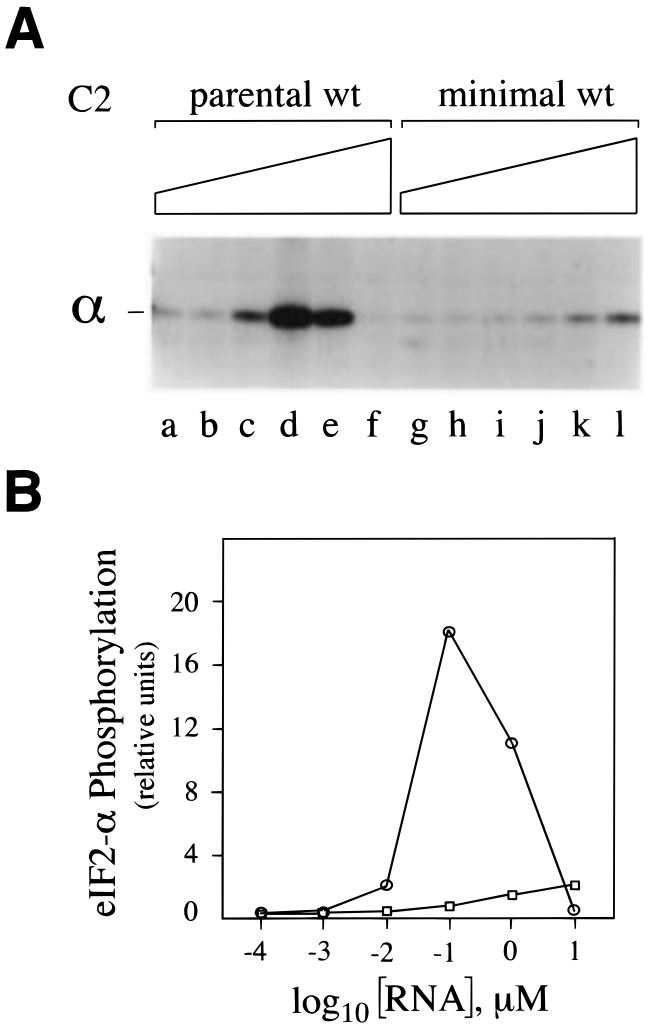

FIGURE 9: In vitro activation of PKR by minimal-length RNAs. (A) Shown is an autoradiogram of a 10% SDS−polyacrylamide gel corresponding to the region of the α subunit of eIF-2. Parental clone 2 wt RNA (lanes a−f) and minimal-length clone 2 wt RNA samples (lanes g−l) were compared for their ability to activate purified PKR kinase as measured by eIF-2α phosphorylation. RNA was added as indicated to the standard reaction mixture which contained [γ-^{32}P]ATP and eIF-2. Concentrations of RNAs used were lanes a and g, 0.0001 μM; lanes b and h, 0.001 μM; lanes c and i, 0.01 μM; lanes d and j, 0.1 μM; lanes e and k, 1.0 μM; lanes f and l, 10 μM RNA. α denotes the position of ^{32}P-labeled α-subunit of eIF-2. (B) Quantitation of autoradiogram shown in panel A (○), full-length clone 2 RNA, (□) minimal-length clone 2 RNA.

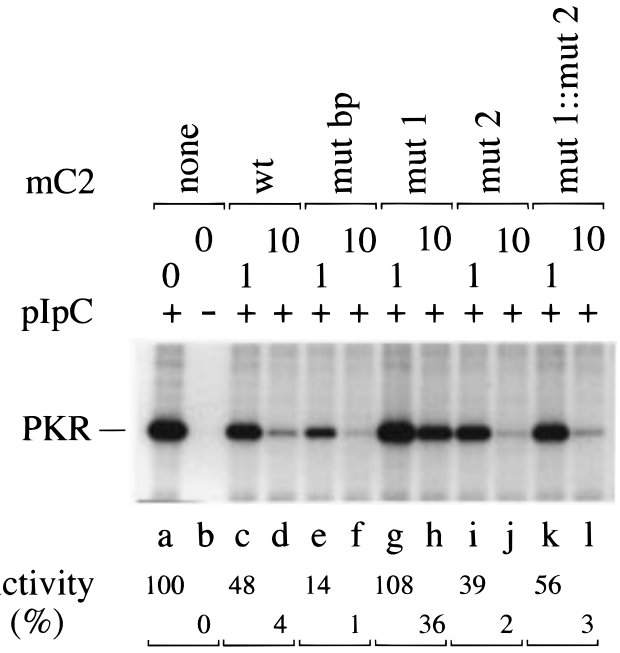

FIGURE 10: Antagonism of PKR activation by minimal-length RNAs. Shown is an autoradiogram of a 10% SDS−polyacrylamide gel. Minimal-length clone 2 (mC2) RNA either wt (lanes c and d) or mutant (lanes e−l), was examined for its ability to antagonize the activation of purified PKR kinase. PKR was incubated with selected RNAs at either 1.0 or 10 μM prior to the addition of 0.1 μg/mL (=0.8 nM) poly(rI)-poly(rC). Reactions carried out either in the absence of added RNA (lane b), or in the presence of the optimal concentration of poly(rI)-poly(rC) (lane a) but without added selected RNA, are shown as controls. PKR denotes the position of ^{32}P-labeled PKR. The autophosphorylation activity of PKR kinase was quantified and is expressed as a percent of the activity observed in the presence of an optimal concentration of poly(rI)-poly(rC).

RNAs selected by binding to the truncated H_6Tp20 PKR protein that consists of residues 1−184 also were able to activate the wild-type 551 residue PKR protein purified from human cells (Figures 3 and 9). However, binding and activation were not equivalent. For example, even though the five RNAs of family 1 shared a 17-nt consensus sequence and displayed similar K_d values with H_6Tp20 protein, they differed significantly in their ability to activate PKR autophosphorylation and eIF-2 phosphorylation. These differences in enzyme activation observed for selected family 1 RNAs may relate to differences in their ability to induce the appropriate conformational change in the full-length protein required for enzymatic activity. While dimerization of full-length PKR can occur both in the presence and absence of RNA binding activity (*64*−*66*), the presence of RNA appears to mediate a conformational change in the catalytic domain of the PKR protein required for activation of kinase activity (*67*, *68*). Furthermore, both genetic (*69*) and biophysical (*68*) analyses suggest that the active form of PKR most likely is a dimer together with a single RNA molecule.

The double-mutant, base-pair mutant, and mutant 2 minimal-length clone 2 RNAs all antagonized the dsRNA-mediated activation of the full-length PKR kinase autophosphorylation as did the wt minimal-length clone 2 RNA (Figure 10). Somewhat surprisingly, however, binding of the mutant 2 RNA to H_6Tp20 protein was not detected by a native-gel mobility-shift assay, whereas the double-mutant, base-pair mutant and wt minimal-length clone 2 RNAs did bind (Figure 7B). This apparent contradiction may reflect the differing sensitivity of the two assays, or alternatively may indicate a transient RNA−protein interaction not detected by the gel-shift assay.

The dsRBD Can Tolerate Non-Watson−Crick Structures. Secondary structures for full-length RNAs from clones 2 and 11 were established by sequence alignment, free-energy minimization, structure mapping, boundary experiments, site-directed mutagenesis, and footprinting experiments. These data are most consistent with the secondary structures shown (Figure 4). These structures include a 3′-hairpin (P3/L3) with a stem that consists of 14 bp counting Watson−Crick, G−U wobble pairs, and the tandem A−G mismatch pair, and a shorter 5′-stem-loop (P2/L2) that is 8 bp in length and is interrupted by a C−U mismatch in one case.

Of these two stem-loop structures, the 3′-most stem loop appears to be the most important for protein binding as judged by the following group of experiments. Boundary experiments probing the 3′-end of the RNA reveal that only the four 3′-most single-stranded nucleotides can be removed from the RNA without affecting binding and that the 3′-stem itself is absolutely required for binding. Boundary experiments examining the 5′-end length requirements of the RNA show that while the 3′-stem-loop is not itself sufficient to bind H_6Tp20, the 5′-end requirements are poorly defined. Inspection of predicted foldings for various 5′-end truncations reveals that they can form minihelices 3−5 bp in length. It is plausible that a minihelix and the noncontiguous 14 bp 3′-stem-loop combine to provide the minimal-length binding site of 16 bp (*22*). This was an unexpected result and suggests that, in general, binding sites for PKR protein may be assembled from several noncontiguous helices. Recent secondary and tertiary structure models for VA I RNA depict two noncontiguous stem loops as being involved in binding PKR, termed the apical stem and stem 7 (*36*). Our results are consistent with such a binding site.

Additionally, for clone 2 RNA, the only observed footprint was in the 3′-stem-loop. For clone 11 RNA, the 3′-stem-loop was protected at 12 different positions while the 5′-stem-loop was protected at only two positions. Interestingly, five of the 12 3′-stem-loop protections occur in the loop itself with one enhancement adjacent to the loop (Figure 4). Protections of the loop may arise from occlusion of the loop by bound H_6Tp20, or perhaps binding of H_6Tp20 to the loop. Enhancement of RNase V1 cleavage may arise from conformational changes of the RNA upon protein binding that render the site optimal for cleavage. The pattern of protections and enhancement observed in the 3′-stem-loop of clone 11 RNA suggests that H_6Tp20 may interact differently with the loop-proximal and loop-distal portions of the stem.

An unexpected feature of the 3′-stem-loop is the presence of a non-Watson−Crick secondary structural element, the tandem A−G mismatch. The role of the tandem A−G mismatch in protein binding was investigated by construction of minimal RNAs containing the 3′-stem-loop plus two extra base pairs to provide a 16 bp binding site. Converting the tandem A−G mismatch to Watson−Crick base pairs had no significant effect on binding affinity, suggesting that the tandem A−G mismatch is merely tolerated within the flanking dsRNA stems. Its selective advantage may then be manifest at some level other than protein binding such as the reverse transcription or PCR amplification step (see below). A separate question is whether the tandem A−G mismatch makes specific contacts with the protein and confers site specificity, as suggested for other secondary structure defects in RNAs (*35*).

The tandem A−G mismatch has been investigated both functionally and structurally by Wu et al. (*56*, *70*). Tandem A−G mismatches constitute thermodynamically stable non-Watson−Crick motifs, stabilizing a helix by as much as −1.2 kcal/mol depending on closing base pairs (*70*). NMR investigations reveal that this stabilization arises because of a stacked, imino-hydrogen-bonded conformation with two hydrogen bonds per A−G mismatch (*56*). The tandem A−G mismatch perturbs the structure in only a minor fashion. The backbone is expanded at the mismatch and the helical rise is decreased; however, an overall A-form geometry is maintained. In addition, these minor changes are propagated only to the neighboring base pairs (*56*). Weak cleavage of the tandem A−G mismatch in clone 11 RNA by RNase V1 (Figures 4 and 5) is consistent with the stacked A−G structure and the known property of RNase V1 to cleave many non-Watson−Crick stacked structures including wobble G−U pairs (*60*). In contrast to the tandem A−G mismatch, tandem G−A mismatches can form a sheared conformation and introduce severe distortion into the backbone (*71*).

Several reports support the regulation of PKR by non-double-stranded RNAs. Early studies showed that PKR could not be activated if a poly(C):poly(I) duplex were interrupted with a G−I mismatch every 7 bp, on average, but could be partially or completely activated if the G−I mismatch were present only every 15 or 45 bp on average, respectively (*72*). PKR can also be activated by reovirus s1 mRNA (*25*) and the 3′-UTR from human α-tropomyosin mRNA (*33*) and inhibited by the adenovirus VA I RNA, Epstein−Barr virus EBER RNAs, and HIV-1 TAR RNA (*32*, *73*), all of which contain secondary structure defects. RNase T1 protection assays reveal that PKR can bind to the genomic RNA from hepatitis delta agent between nucleotides 710−872 (*35*). The δ agent genomic RNA does not contain a tandem A−G mismatch in this stretch of residues, but does contain an internal 5′AA3′/5′AG3′ mismatch (*35*). This latter mismatch is found in clone 14 RNA, the member of family 1 with an exception to the tandem A−G mismatch (Figure 2). It is thus possible that the two mismatch types form similar, dsRNA-like structures. Wu et al. (*70*) surveyed 24 small-subunit and 51 large-subunit prokaryotic and eukaryotic phylogenetic rRNA secondary structures for tandem mismatches. The tandem G−A and the 5′AA3′/5′GA3′ mismatches were found 71 and 54 times, respectively, while the thermodynamically stable tandem A−G mismatch and the 5′AA3′/5′AG3 were not found (*70*). One possible explanation for the absence of these mismatches is that they can regulate dsRBM-containing proteins and as such are not prevalent motifs. Double-stranded RNA deaminases contain multiple copies of the dsRBM and are able to convert A−U pairs to I−U pairs by a deamination reaction (*74*−*77*). Multiple A−U to I−U conversions can occur within the same helix (*27*). Inspection of the I−U pair geometry reveals that it can form a two imino-hydrogen-bonded wobble pair. Apparently, the dsRBM can also bind to double-stranded RNA containing a limited number of I−U wobble pairs.

Design of the experiments in this study may be biased against the selection of long, purely double-stranded RNA regions. Such regions require reverse transcriptase to read through highly structured RNAs. Although the reverse transcription step was performed at 60 °C, such structures were probably selected against. The goal of this study was identification of non-Watson−Crick motifs; therefore, bias against long, regular double-stranded RNAs may have aided the outcome.

Double-stranded regions were found to form between the random region and the fixed primer-binding sites, with the longest duplexes present at the 3′-end. Such pairing may be biased by experimental design. The 3′-primer-binding site is bound first by the bottom-strand primer during reverse transcription. Both primer-binding sites are bound during PCR. Primer may first bind to accessible single-stranded regions adjacent to the RNA stem loop and then invade and

disrupt the neighboring RNA stem loop affording replication. Such disruptions, readily achieved at the 3′-end of the molecule during reverse transcription, are not possible for stems formed within the randomized region only. Since not all non-Watson−Crick possibilities are present in the fixed-sequence primer-binding sites, an exhaustive study of the role of non-Watson−Crick structures in dsRBM binding has not been carried out. Nevertheless, several important conclusion can be made: (1) the binding site for the dsRBD has been found to be composed of RNAs that are similar to, or mimic, A-form dsRNA, one expected outcome based on numerous previous studies; (2) the tandem A−G mismatch can support protein binding; (3) noncontiguous helices can combine to form a site for protein binding; and (4) an experimental basis has been established which will aid design of further selection experiments to examine the role of RNA structure in dsRBD binding.

ACKNOWLEDGMENT

We thank Anne Gooding and Elaine Podell for oligonucleotide synthesis.

REFERENCES

1. Samuel, C. E. (1993) *J. Biol. Chem. 268*, 7603−7606.
2. Clemens, M. J. Protein Kinases That Phosphorylate eIF2 and eIF2B, and Their Role in Eukaryotic Cell Translational Control. in *Translational Control* (Hershey, J. W. B., Mathews, M. B., and Sonenberg, N., Eds.) pp 139−172, Cold Spring Harbor Laboratory, Plainview, NY.
3. Kumar, A., Yang, Y. L., Flati, V., Der, S., Kadereit, S., Deb, A., Haque, J., Reis, L., Weissmann, C., and Williams, B. R. (1997) *EMBO J. 16*, 406−416.
4. Wong, A. H., Tam, N. W., Yang, Y. L., Cuddihy, A. R., Li, S., Kirchhoff, S., Hauser, H., Decker, T., and Koromilas, A. E. (1997) *EMBO J. 16*, 1291−1304.
5. Roy, S., Katze, M. G., Parkin, N. T., Edery, I., Hovanessian, A. G., and Sonenberg, N. (1990) *Science 247*, 1216−1219.
6. Benkirane, M., Neuveut, C., Chun, R. F., Smith, S. M., Samuel, C. E., Gatignol, A., and Jeang, K. T. (1997) *EMBO J. 16*, 611−624.
7. Nagai, K., Wong, A. H., Li, S., Tam, W. N., Cuddihy, A. R., Sonenberg, N., Mathews, M. B., Hiscott, J., Wainberg, M. A., and Koromilas, A. E. (1997) *J. Virol. 71*, 1718−1725.
8. Kitajewski, J., Schneider, R. J., Safer, B., Munemitsu, S. M., Samuel, C. E., Thimmappaya, B., and Shenk, T. (1986) *Cell 45*, 195−200.
9. Mathews, M. B., and Shenk, T. (1991) *J. Virol. 65*, 5657−5662.
10. Lee, S. B., and Esteban, M. (1994) *Virology 199*, 491−496.
11. Der, S. D., Yang, Y. L., Weissmann, C., and Williams, B. R. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 3279−3283.
12. Kibler, K. V., Shors, T., Perkins, K. B., Zeman, C. C., Banaszak, M. P., Biesterfeldt, J., Langland, J. O., and Jacobs, B. L. (1997) *J. Virol. 71*, 1992−2003.
13. Lee, S. B., Rodriguez, D., Rodriguez, J. R., and Esteban, M. (1997) *Virology 231*, 81−88.
14. Katze, M. G., Wambach, M., Wong, M.-L., Garfinkel, M., Meurs, E., Chong, K., Williams, B. R. G., Hovanessian, A. G., and Barber, G. N. (1991) *Mol. Cell. Biol. 11*, 5497−5505.
15. Feng, G.-S., Chong, K., Kumar, A., and Williams, B. R. G. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 5447−5451.
16. Green, S. R., and Mathews, M. B. (1992) *Genes Dev. 6*, 2478−2490.
17. McCormack, S. J., Thomis, D. C., and Samuel, C. E. (1992) *Virology 188*, 47−56.
18. Patel, R. C., and Sen, G. C. (1992) *J. Biol. Chem. 267*, 7671−7676.
19. St Johnston, D., Brown, N. H., Gall, J. G., and Jantsch, M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10979−10983.
20. Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M., and Pastore, A. (1995) *EMBO J. 14*, 3572−3584.
21. Bycroft, M., Grünert, S., Murzin, A. G., Proctor, M., and St. Johnston, D. (1995) *EMBO J. 14*, 3563−3571.
22. Bevilacqua, P. C., and Cech, T. R. (1996) *Biochemistry 35*, 9983−9994.
23. Hunter, T., Hunt, T., Jackson, R. J., and Robertson, H. D. (1975) *J. Biol. Chem. 250*, 409−417.
24. Samuel, C. E. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 600−604.
25. Bischoff, J. R., and Samuel, C. E. (1989) *Virology 172*, 106−115.
26. Manche, L., Green, S. R., Schmedt, C., and Mathews, M. B. (1992) *Mol. Cell. Biol. 12*, 5238−5248.
27. Polson, A. G., and Bass, B. L. (1994) *EMBO J. 13*, 5701−5711.
28. Schweisguth, D. C., Chelladurai, B. S., Nicholson, A. W., and Moore, P. B. (1994) *Nucleic Acids Res. 22*, 604−612.
29. Katze, M. G. (1995) *Trends Microbiol. 3*, 75−78.
30. Proud, C. G. (1995) *Trends Biochem. Sci. 20*, 241−246.
31. Jacobs, B. L., and Langland, J. O. (1996) *Virology 219*, 339−349.
32. Robertson, H. D., and Mathews, M. B. (1996) *Biochimie 78*, 909−914.
33. Davis, S., and Watson, J. C. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 508−513.
34. Robertson, H. D., Manche, L., and Mathews, M. B. (1996) *J. Virol. 70*, 5611−5617.
35. Circle, D. A., Neel, O. D., Robertson, H. D., Clarke, P. A., and Mathews, M. B. (1997) *RNA 3*, 438−448.
36. Ma, Y., and Mathews, M. B. (1996) *RNA 2*, 937−951.
37. Clarke, P. A., Sharp, N. A., and Clemens, M. J. (1990) *Eur. J. Biochem. 193*, 635−641.
38. Sharp, T. V., Schwemmle, M., Jeffrey, I., Laing, K., Mellor, H., Proud, C. G., Hilse, K., and Clemens, M. J. (1993) *Nucleic Acids Res 21*, 4483−4490.
39. Gunnery, S., Rice, A. P., Robertson, H. D., and Mathews, M. B. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8687−8691.
40. Roy, S., Agy, M., Hovanessian, A. G., Sonenberg, N., and Katze, M. (1991) *J. Virol. 65*, 632−640.
41. Maitra, R. K., McMillan, N. A. J., Desai, S., McSwiggen, J., Hovanessian, A. G., Sen, G., Williams, B. R. G., and Silverman, R. H. (1994) *Virology 204*, 823−827.
42. Zhu, S., Romano, P. R., and Wek, R. C. (1997) *J. Biol. Chem. 272*, 14434−14441.
43. Thomis, D. C., Doohan, J. P., and Samuel, C. E. (1992) *Virology 188*, 33−46.
44. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
45. Bartel, D. P., and Szostak, J. W. (1993) *Science 261*, 1411−1418.
46. Czerny, T. (1996) *Nucleic Acids Res. 24*, 985−986.
47. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
48. Schmedt, C., Green, S. R., Manche, L., Taylor, D. R., Ma, Y., and Mathews, M. B. (1995) *J. Mol. Biol. 249*, 29−44.
49. Zaug, A. J., Dávila-Aponte, J. A., and Cech, T. R. (1994) *Biochemistry 33*, 14935−14947.
50. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
51. Vlassov, V. V., Zuber, G., Felden, B., Behr, J.-P., and Giege, R. (1995) *Nucleic Acids Res. 23*, 3161−3167.
52. Donis-Keller, H., Maxam, A. M., and Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527−2538.
53. Samuel, C. E., Knutson, G. S., Berry, M. J., Atwater, J. A., and Lasky, S. R. (1986) *Methods Enzymol. 119*, 499−516.
54. Thomis, D. C., and Samuel, C. E. (1995) *J. Virol. 69*, 5195−5198.
55. Doudna, J. A., Cech, T. R., and Sullenger, B. A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 2355−2359.
56. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) *Biochemistry 36*, 4449−4460.

57. Jaeger, J. A., Turner, D. H., and Zuker, M. (1990) *Methods Enzymol. 183*, 281−306.
58. Serra, M. J., and Turner, D. H. (1995) *Methods Enzymol. 259*, 242−261.
59. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P., and Ehresmann, B. (1987) *Nucleic Acids Res. 15*, 9109−9128.
60. Lockard, R. E., and Kumar, A. (1981) *Nucleic Acids Res. 9*, 5125−5140.
61. Kolchanov, N. A., Titov, I. I., Vlassova, I. E., and Vlassov, V. V. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 53*, 131−196.
62. Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., and Turner, D. (1985) *Biochemistry 24*, 4533−4539.
63. Dunn, J. J., and Studier, F. W. (1975) *J. Mol. Biol. 99*, 487−499.
64. Patel, R. C., Stanton, P., McMillan, N. M. J., Williams, B. R. G., and Sen, G. C. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8283−8287.
65. Ortega, L. G., McCotter, M. D., Henry, G. L., McCormack, S. J., Thomis, D. C., and Samuel, C. E. (1996) *Virology 215*, 31−39.
66. Wu, S., and Kaufman, R. J. (1997) *J. Biol. Chem. 272*, 1291−1296.
67. Bischoff, J. R., and Samuel, C. E. (1985) *J. Biol. Chem. 260*, 8237−8239.
68. Carpick, B. W., Graziano, V., Schneider, D., Maitra, R. K., Lee, X., and Williams, B. R. G. (1997) *J. Biol. Chem. 272*, 9510−9516.
69. Romano, P. R., Green, S. R., Barber, G. N., Mathews, M. B., and Hinnebusch, A. G. (1995) *Mol. Cell. Biol. 15*, 365−378.
70. Wu, M., McDowell, J. A., and Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.
71. SantaLucia, J., Jr., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
72. Minks, M. A., West, D. K., Benvin, S., and Baglioni, C. (1979) *J. Biol. Chem. 254*, 10180−10183.
73. Clemens, M. J., Laing, K. G., Jeffrey, I. W., Schofield, A., Sharp, T. V., Elia, A., Matys, V., James, M. C., and Tilleray, V. J. (1994) *Biochimie 76*, 770−778.
74. Kim, U., Wang, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11457−11461.
75. O'Connell, M. A., Krause, S., Higuchi, M., Hsuan, J. J., Totty, N. F. T., Jenny, A., and Keller, W. (1995) *Mol. Cell. Biol. 15*, 1389−1397.
76. Patterson, J. B., and Samuel, C. E. (1995) *Mol. Cell. Biol. 15*, 5376−5388.
77. Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P. H., and Higuchi, M. (1996) *Nature 379*, 460−464.

BI980113J