Specificities of Caenorhabditis elegans and human hairpin binding proteins for the first nucleotide in the histone mRNA hairpin loop

F. Michel, D. Schumperli and B. Muller

RNA 2000 6: 1539-1550

RNA (2000), *6*:1539–1550. Cambridge University Press. Printed in the USA.

Specificities of *Caenorhabditis elegans* and human hairpin binding proteins for the first nucleotide in the histone mRNA hairpin loop

FABRICE MICHEL,[1,3] DANIEL SCHÜMPERLI,[1] and BERNDT MÜLLER[1,2]

[1]Institute of Cell Biology, University of Bern, 3012 Bern, Switzerland
[2]Department of Molecular and Cell Biology, Institute of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, Scotland, UK

ABSTRACT

The 3′ ends of animal replication-dependent histone mRNAs are formed by endonucleolytic cleavage of the primary transcripts downstream of a highly conserved RNA hairpin. The hairpin-binding protein (HBP) binds to this RNA element and is involved in histone RNA 3′ processing. A minimal RNA-binding domain (RBD) of ~73 amino acids that has no similarity with other known RNA-binding motifs was identified in human HBP [Wang Z-F et al., *Genes & Dev,* 1996, *10*:3028-3040]. The primary sequence identity between human and *Caenorhabditis elegans* RBDs is 55% compared to 38% for the full-length proteins. We analyzed whether differences between *C. elegans* and human HBP and hairpins are reflected in the specificity of RNA binding. The *C. elegans* HBP and its RBD recognize only their cognate RNA hairpins, whereas the human HBP or RBD can bind both the mammalian and the *C. elegans* hairpins. This selectivity of *C. elegans* HBP is mostly mediated by the first nucleotide in the loop, which is C in *C. elegans* and U in all other metazoans. By converting amino acids in the human RBD to the corresponding *C. elegans* residues at places where the latter deviates from the consensus, we could identify two amino acid segments that contribute to selectivity for the first nucleotide of the hairpin loop.

Keywords: RNA-binding domain; RNA–protein interaction; stem-loop-binding protein

INTRODUCTION

Replication-dependent histone mRNAs are not polyadenylated, but are formed at the 3′ end by an endonucleolytic cleavage of the primary transcripts. This nuclear cleavage event requires two *cis*-acting elements: a purine-rich sequence (or spacer element) downstream of the cleavage site that is recognized by the U7 small nuclear ribonucleoprotein (snRNP) through RNA–RNA base pairing (Mowry & Steitz, 1987a; Cotten et al., 1988; Soldati & Schümperli, 1988; Bond et al., 1991) and a conserved 26-nt sequence containing a hairpin that interacts with a factor called hairpin-binding protein (HBP) or stem-loop-binding protein (SLBP) (Mowry & Steitz, 1987b; Vasserot et al., 1989; Melin et al., 1992; Wang et al., 1996; Martin et al., 1997). The cleavage of histone pre-mRNA requires a third factor, termed heat-labile factor (HLF) that has not been well characterized (Gick et al., 1987). Besides histone RNA 3′ processing, the hairpin is also essential for nuclear export, translation, and stability regulation of histone mRNAs (reviewed in Marzluff, 1992; Marzluff & Hanson, 1993; Wittop Koning & Schümperli, 1994; Müller & Schümperli, 1997). As HBP is also present on polysome-associated histone mRNA (Pandey et al., 1991; Dominski et al., 1995), it is thought to control most, if not all, posttranscriptional aspects of histone gene expression and regulation.

The sequence of the hairpin element in vertebrate and *C. elegans* histone genes varies both in the loop and in the sequences flanking the hairpin. One difference in the loop is the nucleotide at position 1 of the U-rich 4-nt loop. In all other species, including humans, there is a U at this position, whereas in *C. elegans* it is a C (Marzluff, 1992; Marzluff & Hanson, 1993; Wittop Koning & Schümperli, 1994). The situation is reversed at the 3′ end of the loop where the *C. elegans* hairpin has a U, whereas in other species, this nucleotide is less well conserved but most often a C.

Human, mouse, *Xenopus,* and *C. elegans* HBPs have been identified that bind specifically to this histone RNA

Reprint requests to: Daniel Schümperli, Institut für Zellbiologie, Universität Bern, Baltzerstrasse 4, 3012 Bern, Switzerland; e-mail: daniel.schuemperli@izb.unibe.ch.
[3]Present address: EMBL-Grenoble Outstation, Protein-DNA interaction group, c/o ILL BP 156, 38042 Grenoble Cedex 9, France.

hairpin element (Wang et al., 1996, 1999; Martin et al., 1997). An ~73-amino-acid minimal RNA-binding domain (RBD) of HBP has been characterized by deletion analysis (Wang et al., 1996). This RBD has no obvious primary sequence homology with other RNA-binding motifs characterized so far (Burd & Dreyfuss, 1994; Draper, 1995, 1999). The RBD also encompasses a short 67-amino-acid region (residues 9–75 in Fig. 1) that is the most conserved part among the known animal HBPs (Wang et al., 1996, 1999; Martin et al., 1997). In this region, *C. elegans* and human HBP are 55% identical, whereas the overall sequence identity between these two proteins is 38%. This is lower than the ~63% identity in the RBD between the human HBP and *Xenopus* SLBP2, the most distantly related vertebrate HBPs. Here we have determined experimentally whether these differences between *C. elegans* and human HBPs are reflected by and contribute to a selective RNA binding. These experiments revealed that the *C. elegans* HBP binds with great preference to *C. elegans* hairpin RNA, and that the *C. elegans* RBD is responsible for this preference. Human HBP, on the other hand, as well as the human RBD, did not show a strong preference for either vertebrate or *C. elegans* hairpin RNA. Binding reactions with mutant hairpin RNAs demonstrated that the first nucleotide in the loop is an important specificity determinant for the binding of the *C. elegans* RBD to its cognate RNA. To identify amino acid sequence elements involved in this binding specificity, we have changed single or multiple residues in the human RBD to the corresponding *C. elegans* sequence and tested these chimeric proteins for RNA binding. This approach enabled us to identify two amino acid segments of HBP that participate in RNA binding through interaction with specific nucleotides in the hairpin loop.

RESULTS

The RNA-binding domain of *C. elegans* HBP is sufficient to discriminate against binding to mammalian hairpin RNA

To know if the difference between human and *C. elegans* RBD sequences was reflected by different specificities for histone hairpin RNAs, we produced human and *C. elegans* HBP in vitro and tested these proteins in electrophoretic mobility shift assays (EMSA) with ^{32}P-labeled RNAs containing either a mammalian or a *C. elegans* histone mRNA hairpin. The mammalian hairpin illustrated in Figure 2 as mmHPs RNA is derived from the mouse histone H4.12 gene (Meier et al., 1989). The *C. elegans* sequence referred to here as ceHP RNA is from the *C. elegans* H4-3.2 gene (Roberts et al., 1989). The ceHPs RNA, a shorter variant of ceHP RNA, shows the same binding properties as the longer *C. elegans* substrate but has a reduced propensity to dimerize (Fig. 2).

Interestingly, although human HBP bound equally well to the mmHPs and ceHPs RNAs (Fig. 3B, lanes 2 and 5), the *C. elegans* HBP bound almost exclusively to ceHPs RNA (Fig. 3B, lanes 3 and 6). This difference remained when HBP fragments lacking either the C- or N-terminal sequences flanking the RBD were used (Fig. 3A,B). Truncated proteins derived from human HBP formed complexes with both mmHPs (Fig. 3B, lanes 8 and 9) and ceHPs RNA (Fig. 3B, lanes 13 and 14), whereas complexes with protein fragments derived from *C. elegans* HBP were only observed with ceHPs RNA (Fig. 3B, lanes 15 and 16). We observed multiple bands with proteins containing the human N-terminus (Fig. 3B, lanes 2, 5, 8, and 13). Some of these complexes may have arisen either from modified

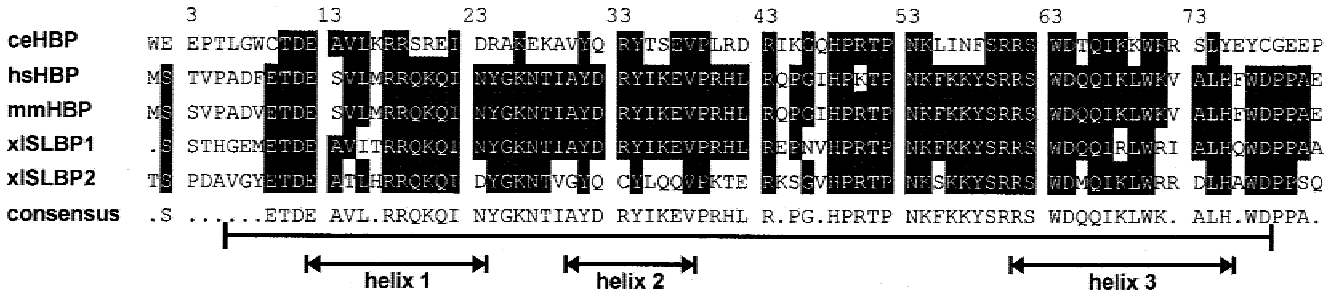

FIGURE 1. Sequence alignment of the minimal RNA-binding domains (RBDs) of histone hairpin binding proteins (HBPs). The numbering of residues indicated at the top reflects the 82 amino acid RBD peptides of human and *C. elegans* HBP used in this study. Amino acids identical in at least three of the five HBPs analyzed are indicated with white text in black boxes and shown again in the consensus sequence. The minimal 73-amino-acid RBD defined for human HBP by Wang et al. (1996) is emphasized by a line below the consensus sequence. Three α-helical regions predicted by the PHDsec program (Rost & Sander, 1993, 1994; see Martin et al., 2000) are indicated by double arrows. ce: *C. elegans*; hs: human; mm: mouse; xl: *X. laevis*. Note that *X. laevis* has two HBP-like proteins—SLBP1, related to the mammalian nuclear RNA processing factor HBP, and SLBP2, involved in storing translationally inactive maternal histone mRNAs in the oocyte (Wang et al., 1999).

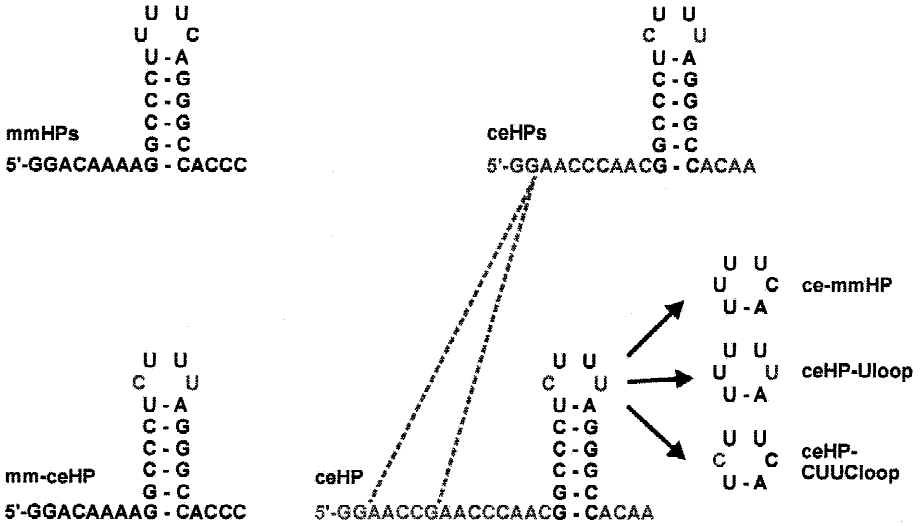

FIGURE 2. Histone hairpin RNAs used in this work. The mammalian (mmHPs) and the *C. elegans* (ceHPs or ceHP) RNA sequences are derived from the mouse H4.12 gene (Meier et al., 1989) and the *C. elegans* H4-3.2 gene (Roberts et al., 1989), respectively. The ceHPs has a shorter flanking sequence than the ceHP RNA and gray dashed lines indicate the deletion. The sequence differences between the ceHP and mmHPs RNAs are shown in gray in the *C. elegans* sequence. Mutant RNAs: The mm-ceHP has the sequences flanking the mammalian hairpin structure and the *C. elegans* hairpin sequence; ce-mmHP has the *C. elegans* flanking sequence and the mammalian hairpin sequence. ceHP-Uloop and ceHP-CUUCloop RNAs have only one mutation at the fourth and first positions of the *C. elegans* hairpin sequence, respectively.

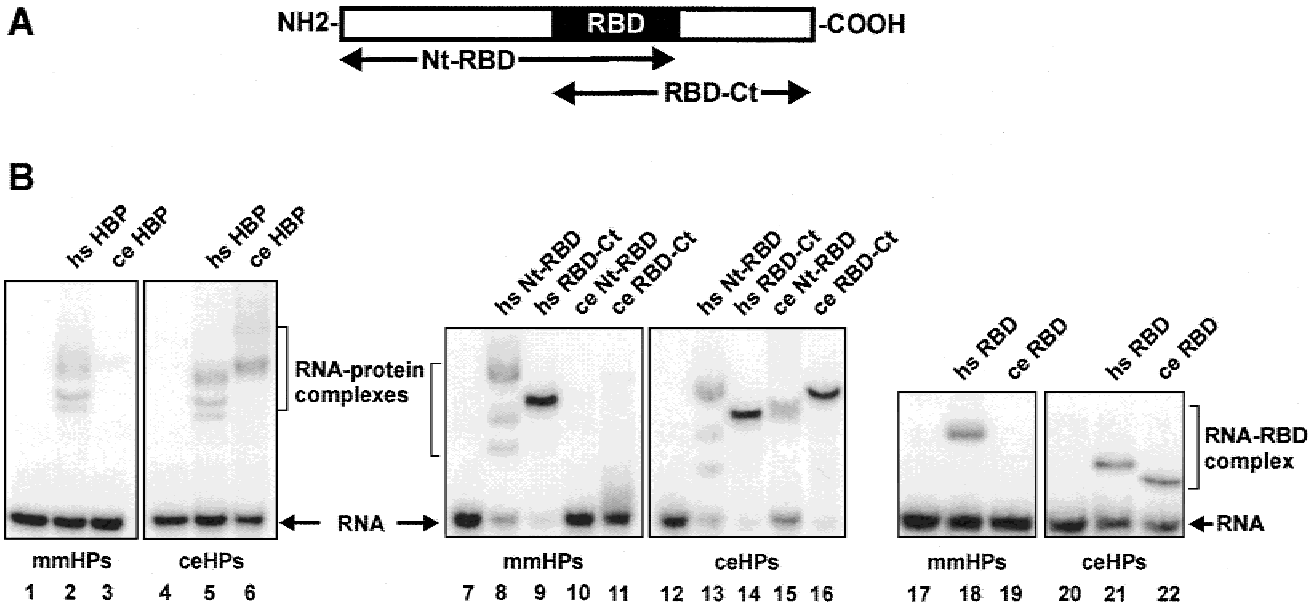

FIGURE 3. RNA-binding domains of human and *C. elegans* HBP bind with similar specificity to the respective full-length proteins to mammalian and *C. elegans* hairpins. **A**: Schematic representation of the hairpin-binding protein. The central black region is the RBD used in this study. Double arrows indicate truncated HBP variants lacking either the C- or the N-terminal sequences flanking the RBD. **B:** In vitro-binding assays. The indicated full-length or truncated HBPs made in vitro were incubated with ^{32}P-labeled mmHPs or ceHPs RNAs as described in Materials and Methods. The protein–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis (EMSA). Free RNAs are shown in lanes 1, 7, and 17 for mmHPs and in lanes 4, 12, and 20 for ceHPs. Additional complexes obtained with the hs HBP (lanes 2 and 5) and hs Nt-RBD (lanes 8 and 13) preparations are most likely the result of modified proteins or minor translation products. The mobility difference between the complexes formed by hs RBD with mmHPs (lane 18) or ceHPs RNA (lane 21) may be due to the different sequences and conformations of the two RNAs and was also observed in other experiments (see Figs. 5 and 6).

proteins (i.e., posttranslational modifications) or from minor translation products (i.e., internal translation initiation), but all of them showed identical binding behavior. We thus concluded that the structural elements responsible for the discrimination of the *C. elegans* HBP against binding to mmHPs RNA must be located within its RBD.

This was confirmed in binding studies with the human HBP fragment encompassing the minimal RBD and the corresponding *C. elegans* protein fragment illustrated in Figure 1. In this article, we refer to these shorter proteins as human (hs) and *C. elegans* (ce) RBD, respectively. As expected, the hs RBD was able to form complexes with both ceHPs and mmHPs RNA (Fig. 3B, lanes 18 and 21), whereas the ce RBD interacted only with its cognate RNA substrate and not with the mmHPs RNA (Fig. 3B, lanes 19 and 22). We then determined dissociation constants (K_d) for the human and *C. elegans* RBDs as described in Materials and Methods and in the legend to Figure 4. The hs RBD bound to both mmHPs and ceHPs RNAs with a K_d of ~4 nM (Fig. 4A and Table 1). The ce RBD displayed a slightly higher affinity for ceHPs RNA ($K_d \approx 3$ nM). However, its binding to mmHPs RNA was very weak (Fig. 4B) and, as no clear-cut saturation was reached within the range of RNA concentrations tested, the K_d for mmHPs RNA could not be determined accurately. Nevertheless, the data indicate that the K_d must be bigger than 7 nM (see Table 1). As is shown in Figure 7, these differences in binding behavior persisted over a wide range of salt concentrations. All these results clearly indicate that the *C. elegans* RBD was sufficient to discriminate against binding to mammalian hairpin RNA.

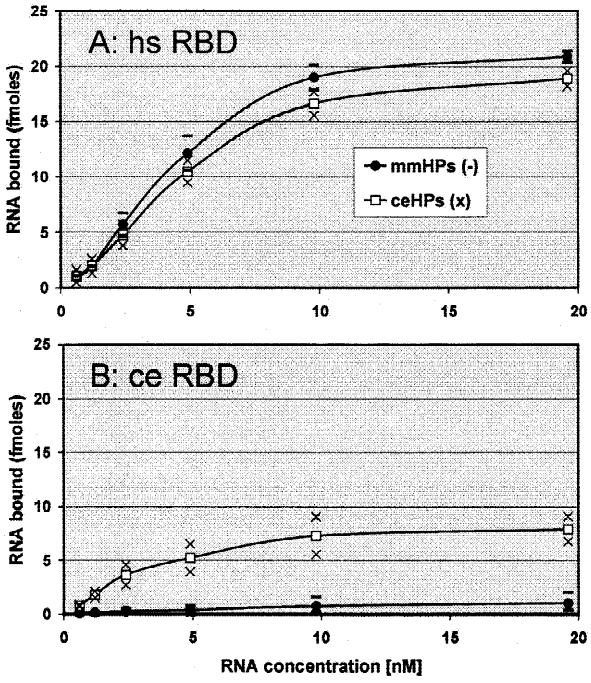

FIGURE 4. Determination of dissociation constants for the interaction of human (**A**) and *C. elegans* (**B**) RBDs with mouse and *C. elegans* RNA hairpins. The indicated in vitro-synthesized RBDs (4.4 nM) were incubated with increasing amounts of ^{32}P-labeled mmHPs or ceHPs RNAs as described in Materials and Methods. The protein–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis (EMSA) and the amount of complexed RNA was determined by PhosphorImager. Plotted is the amount of RNA bound at each RNA concentration versus the RNA concentration, allowing for a direct reading of the dissociation constant K_d (whose values are listed in Table 1). Experiments were done at least in duplicate, and shown are the means, flanked by the highest and lowest values at each RNA concentration. Symbols used for the two RNAs are defined in the insert of **A**.

TABLE 1. Binding of human and *C. elegans* RBDs and human RBD mutants to mmHPs and ceHPs RNAs

		mmHPs[b]		ceHPs[b]	
Protein	Binding ratio[a] (mmHPs/ceHPs)	K_d (nM)	Max binding (fmol)	K_d (nM)	Max binding (fmol)
hs RBD	1.03	4	21	4	19
ce RBD	0.05	(>7)	nd	3	8
mut3	0.50	(>7)	nd	5	22
mut7	0.39	(>8)	nd	6	9
mut9	12	3	2.8	(>7)	nd

[a]The RNA–protein complexes from Figure 6 and from two further identical experiments were quantitated by PhosphorImager and the average ratios of mmHPs complexes to ceHPs complexes are listed.
[b]Dissociation constants and maximal binding activities were read from Figure 4 (hs and ce RBD) and from similar graphs (mut3, mut 7, and mut9). Where no saturation was reached within the range of RNA concentrations tested, a minimal value for K_d is listed in brackets and the maximal binding is given as nd.

The nucleotides in the histone hairpin loop are important for sequence discrimination

To further investigate this difference in binding specificity between the human and *C. elegans* HBPs and RBDs, we tried to identify which features of the RNA hairpins were responsible for this differential recognition. Because the two RBDs behaved similarly to the full-length HBPs, we decided to use them for further experiments. We first produced two chimeric RNA molecules, mm-ceHP and ce-mmHP, where the sequences flanking the hairpin were exchanged (Fig. 2). We used these two RNAs in an EMSA with both *C. elegans* and human RBD. The hs RBD bound both RNAs (Fig. 5A, lanes 5 and 8) in addition to the original mmHPs and ceHP RNAs (Fig. 5A, lanes 2 and 11). In contrast, the ce RBD bound only mm-ceHP RNA and the original ceHP RNA (Fig. 5A, lanes 6 and 12), that is, the RNAs with the *C. elegans* hairpin sequence, but not ce-mmHP or mmHPs RNAs (Fig. 5A, lanes 3 and 9), indicating that the contribution of the flanking sequences to the specificity of binding was only moderate (see quantitation in Fig. 5B).

We therefore concentrated our analysis on the hairpin itself. The hairpin RNA sequence is highly

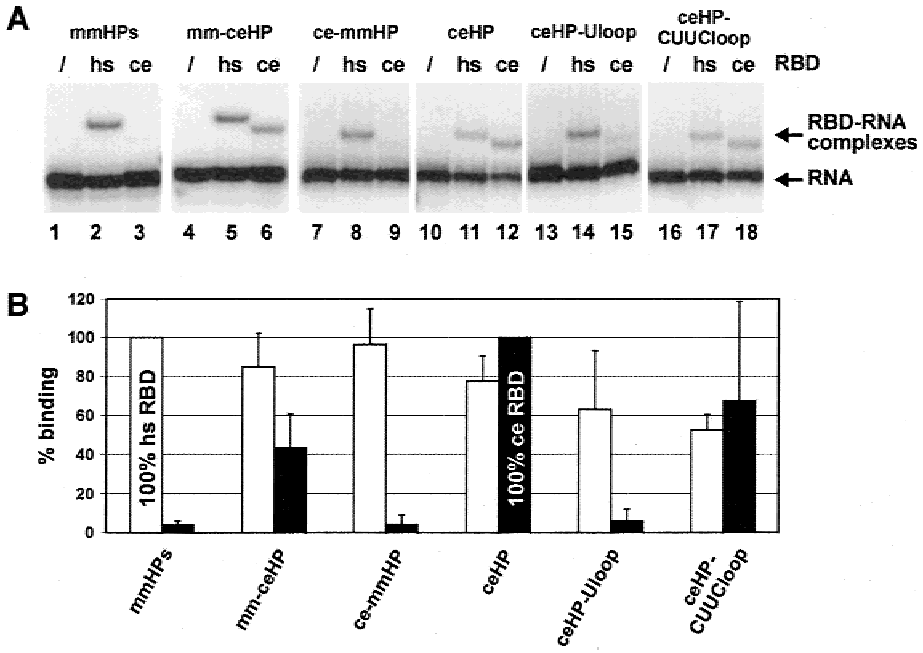

FIGURE 5. Nucleotides in the hairpin loop are important for the high RNA binding selectivity of the *C. elegans* RBD. **A**: Electrophoretic mobility shift assays using the different RNAs described in Figure 2 and the human and *C. elegans* RBDs. The designation at the top of each panel indicates the RNA. The RNAs were incubated without protein (/), with human RBD (hs), or with *C. elegans* RBD (ce) as described in Materials and Methods. The protein–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis (EMSA). **B**: Summary of binding data. The resultant RNA–protein complexes (from **A**) were quantified by PhosphorImager. The average of at least three independent determinations of the relative binding (mmHPs RNA = 100% for the hs RBD; ceHP RNA = 100% for the ce RBD) is represented in this diagram; hs and ce RBDs are shown with white and black bars, respectively; error bars indicate standard deviations.

conserved. One difference between nematode and vertebrate histone hairpins is the nucleotide at position 1 of the U-rich 4-nt loop. In vertebrates, it is invariably a U; in contrast, *C. elegans* is the only exception among all metazoans where a C is found at this first position (Marzluff, 1992; Marzluff & Hanson, 1993; Wittop Koning & Schümperli, 1994). To test whether a U at this position inhibits ce RBD binding, we replaced it by a U in the ceHP sequence, to produce ceHP-Uloop RNA (Fig. 2). Whereas the hs RBD was able to bind to this RNA, complex formation with the ce RBD was greatly reduced to ~5–10% of complexes formed with ceHP RNA (Fig. 5A, lanes 14 and 15; Fig. 5B). This indicated that a C at the 5′ position of the histone hairpin loop is crucial for the binding of the *C. elegans* RBD, whereas the human RBD can bind to the hairpin irrespective of the kind of pyrimidine present at this position.

A second difference between the vertebrate and *C. elegans* hairpins is the nucleotide at the fourth position of the hairpin loop. All *C. elegans* histone genes have a U, whereas vertebrate histone genes can have any nucleotide, but most often have a C at this position. To test whether this nucleotide is important for the sequence discrimination of the *C. elegans* RBD, we replaced it by a C in ceHP RNA, to produce ceHP-CUUCloop RNA (Fig. 2). Both hs and ce RBDs were able to form complexes with this RNA (Fig. 5A, lanes 17 and 18; Fig. 5B), indicating that this residue was not important for sequence discrimination. For the human RBD, ceHP-CUUCloop RNA (Fig. 5A, lane 17) was the least efficient substrate (~75% binding compared to ceHP RNA; Fig. 5B), indicating that replacing the U at position 1 in the loop by a C also slightly affected RNA binding.

Identification of amino acids in the human and *C. elegans* RBD contributing to RNA binding specificity

To identify which features of the RBD are involved in the recognition of the critical U or C nucleotides in the RNA hairpin loop, we tested whether replacing specific amino acids with the corresponding *C. elegans* ones would render the human RBD more selective for *C. elegans* hairpin RNA. Based on the alignment of the five known HBP sequences shown in Figure 1, a consensus sequence of the RBD was derived (Figs. 1 and 6A). All residues shown in the consensus sequence are present in at least three of the five proteins. Of the 32 amino acids of the *C. elegans* RBD that deviate from the consensus, 13 (indicated by gray boxes in Fig. 6A) were investigated. The others were not analyzed, either because at least one of the vertebrate proteins also deviated from the consensus at this

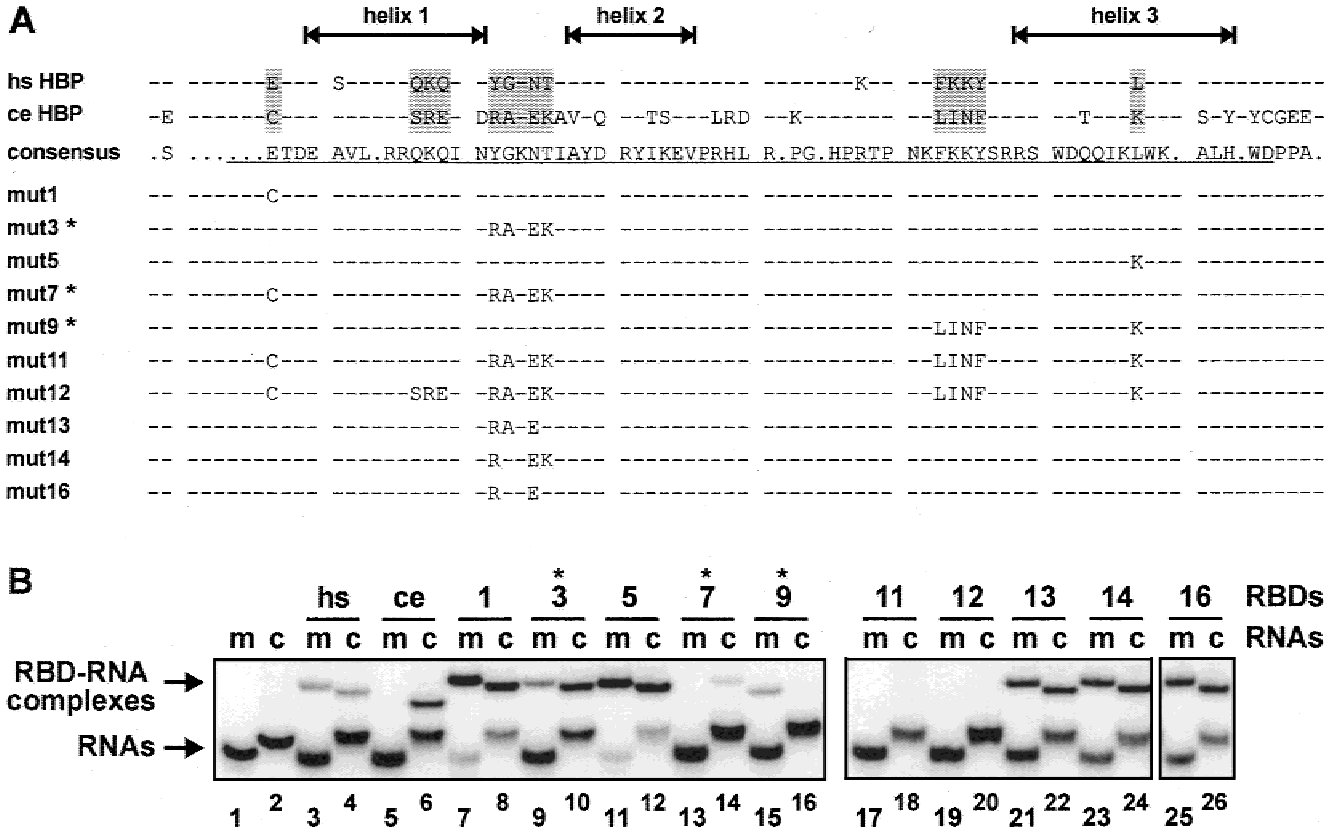

FIGURE 6. Analysis of RNA binding selectivity of human RBD mutants with amino acid sequence replacements making them more similar to the *C. elegans* RBD. **A**: For the human and *C. elegans* RBDs, only the amino acids that differ from the consensus sequence are shown. Gray boxes indicate the amino acids exchanged in this study. The minimal 73-amino-acid RBD defined for human HBP by Wang et al. (1996) is emphasized by underlining of the consensus sequence. Three α-helical regions predicted by the PHDsec program (Rost & Sander, 1993, 1994; see Martin et al., 2000) are indicated by double arrows. For the hs RBD mutants (mut1 to mut16) only the mutated residues are shown. Asterisks indicate the three most interesting mutants described in the text. **B**: Electrophoretic mobility shift assay. The human RBD (hs), the *C. elegans* RBD (ce), and the different mutant RBDs (1–16) were incubated with ^{32}P-labeled mmHPs (m) or ceHPs (c) RNAs as described in Materials and Methods. The protein–RNA complexes were separated from free RNA by EMSA. Free mmHPs and ceHPs RNAs alone are shown in lanes 1 and 2, respectively.

position or because the residues were at the very C-terminal border of the RBD. The amino acids of the human RBD within each of the five gray boxes shown in Figure 6A were replaced with the corresponding *C. elegans* residues using mutagenic oligonucleotide primers in site-directed mutagenesis (see Materials and Methods). Combinations of two or more primers allowed the construction of more complex exchanges. In some cases, additional changes not corresponding to the primer sequence were detected by sequencing. All the resulting proteins (mut1 to mut16, shown in Fig. 6A) were tested for their ability to bind the mmHPs and ceHPs RNA by EMSA (Fig. 6B). Although the in vitro translation of the different polypeptides was monitored by ^{35}S-methionine incorporation followed by SDS-PAGE and autoradiography (data not shown), this does not mean that equal amounts of active proteins were used and hence comparative conclusions between different proteins should not be drawn. However, the main purpose of this experiment was to analyze the relative binding of a given protein to different RNAs, for which it was sufficient that active protein be available. The experiments were done in triplicates and the amount of complexed RNA was determined for each reaction.

As already observed in Figures 3–5, the human RBD bound both RNA molecules with similar efficiency (Fig. 6B, lanes 3 and 4). The average ratio between the amount of complex formed with mmHPs RNA and that formed with ceHPs RNA was 1.03 (Table 1). The *C. elegans* RBD that bound exclusively to its own RNA (Fig. 6B, lanes 5 and 6) had a corresponding average ratio of 0.05 (Table 1).

Mutant RBDs where all the sequence elements were exchanged to make them most similar to the *C. elegans* RBD (mut12), or where all except QKQ (amino acids 19–21) were exchanged (mut11), were not able to bind either RNA (Fig. 6B, lanes 17–20), although both proteins were added in significant amounts (data not shown). This was surprising because we expected these peptides to behave similarly to the *C. elegans*

RBD, and it suggested that in these mutants we had disrupted essential contacts or destabilized the structure of the protein.

Two single amino acid replacements, in mut1 (Fig. 6B, lanes 7 and 8) and mut5 (lanes 11 and 12), did not significantly alter the binding behavior compared to the hs RBD, as the ratios of complex formation for these mutants were 1.16 and 1.14, respectively.

In contrast, mut7 (replacing the sequence YGKNT (amino acids 24–28) by RAKEK in combination with E9C) had almost lost the ability to bind to mmHPs RNA (Fig. 6B, lanes 13 and 14; average ratio of complex formation 0.39, Table 1). Replacement of only YGKNT (24-28) by RAKEK resulted in a protein (mut3) with increased binding to ceHPs or reduced binding to mmHPs RNA but that still bound both RNAs (Fig. 6B, lanes 9 and 10; ratio of 0.5, Table 1). Both mutants point to the sequence YGKNT(24–28) as being important for binding specificity. Exchanging amino acids YGKN(24–27) by RAKE (mut13), or by RGKE (mut16) resulted in proteins that exhibited a similar but weaker shift in selectivity, as these two mutants showed a slight preference for binding to ceHPs RNA (Fig. 6B, lanes 21–22 and 25–26; ratios of 0.77 and 0.85, respectively). Only mut14, where the same sequence was changed to RGKEK, had a binding behavior almost indistinguishable from the hs RBD (Fig. 6B, lanes 23 and 24; average ratio of 1.08).

Interestingly, replacing the sequence FKKY (amino acids 55–58) with LINF (mut9) strongly increased the preference of the RBD for the mmHPs RNA (Fig. 6B, lanes 15 and 16) to the extent that the ratio of complex formation increased to the average value of 12 (Table 1). This result indicates that by replacing this sequence, we have created an RBD with an increased selectivity for mammalian hairpin RNA.

We next determined dissociation constants for the complexes formed between the mut3, mut7, and mut9 RBDs and mmHPs and ceHPs RNAs. The methods used were identical to the ones described for the hs and ce RBDs in Figure 4. The mut9 RBD bound to mmHPs with a similar or slightly stronger affinity as hs RBD, but the maximal binding was reduced to ~2.8 fmol (Table 1). The binding to ceHPs was about 3.5-fold weaker averaged over the entire range of concentrations tested (0.9–19.6 nM) and, because no saturation was reached, the K_d could not be determined precisely (Table 1). This confirmed that mut9 bound with a marked preference for mmHPs over ceHPs RNA. For the mutant RBDs mut3 and mut7, the affinity for the ceHPs RNA was slightly lower (~5 and 6 nM, respectively) than that of either hs RBD or ce RBD for the same substrate. The maximal binding of mut3 and mut7 to ceHPs RNA was comparable to that of hs RBD and ce RBD, respectively. However the binding of both mutant RBDs to mmHPs RNA compared to the binding to ceHPs RNA was reduced at all concentrations (approximately twofold for mut3 and threefold for mut7) and saturation was not reached with these proteins. This again resulted in a difficulty in determining binding constants (Table 1), but confirmed the earlier results presented in Figure 6 that these mutants bound preferentially to ceHPs RNA. For all those combinations where the data did not allow a precise determination of K_d values, the lowest possible values based on these experiments are listed in Table 1.

Finally, we tested the possibility that the observed differences in binding could be dependent on the particular salt concentrations used. We varied the salt concentration from only the amount carried over from the in vitro translation system (~15 mM potassium acetate) up to 1,000 mM KCl. Complex formation showed only a slight increase towards the lower salt concentrations, but was virtually constant between 300 and 1,000 mM salt (Fig. 7). Most importantly, the relative binding behavior of the different RBDs did not change over the range of salt concentrations tested. In this

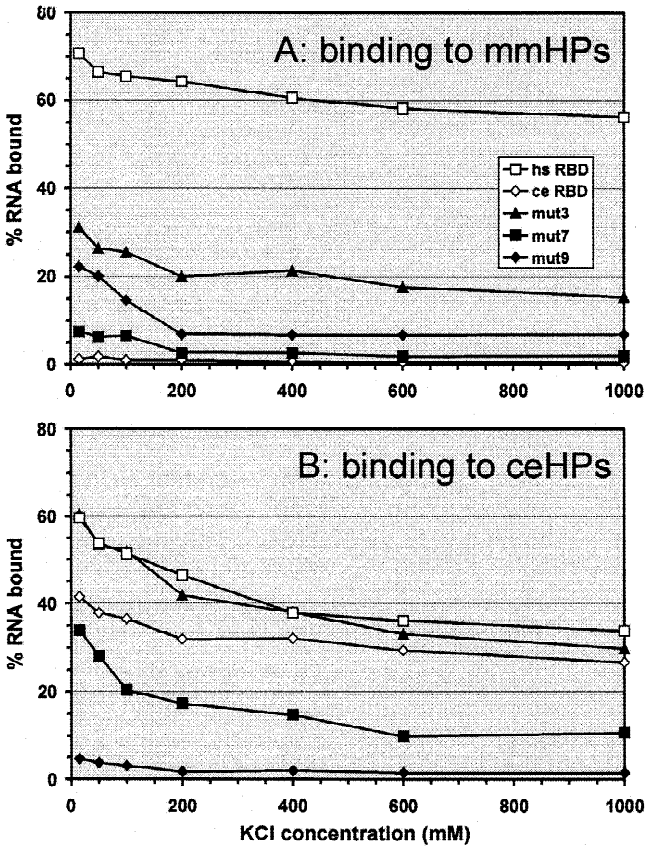

FIGURE 7. RNA complex formation of different RBDs at varying salt concentrations. The in vitro-synthesized human, *C. elegans*, and mut3, mut7, and mut9 RBDs (4.4 nM) were incubated with 4.9 nM ^{32}P-labeled mmHPs (**A**) or ceHPs (**B**) RNAs either without salt (15 mM value) or in 50 mM KCl (see Materials and Methods). Subsequently, the samples were supplemented with KCl to yield the indicated concentrations and further incubated for 20 min on ice prior to analysis by EMSA. Plotted are the binding efficiencies (in percentage of RNA bound) against the different KCl concentrations. Symbols used for the different RBDs are defined in the insert of **A**.

experiment, the hs RBD bound slightly better to mmHPs than to ceHPs (binding ratio averaged over all salt concentrations = 1.42). For mut9, the binding ratio was 4.44, that is, the protein showed a strong preference for binding to mmHPs RNA at all salt concentrations tested. The other RBDs preferentially bound to ceHPs RNA with the preference decreasing in the order ce RBD, mut7, mut3 (binding ratios of 0.03, 0.21, and 0.50, respectively). The absolute levels of binding observed in this experiment were consistent with the results of the RNA titrations presented in Figure 4 and Table 1.

The above results identified two sequence elements in the human RBD that are important for selective RNA binding, the first element YGKNT(24–28) that is changed in mut3 and mut7 and the second one FKKY(55–58) that is replaced by the corresponding *C. elegans* sequence in mut9. We therefore wanted to know whether the new discrimination between *C. elegans* and vertebrate hairpins displayed by these mutants also involved the first nucleotide in the hairpin loop. The results of EMSA using the mutant RNAs described in Figure 2 are summarized in Figure 8. The binding of mut7 and mut9 proteins to ceHP and mmHPs RNAs, respectively, were defined as 100%. The mut7 protein bound preferentially to all RNAs containing a C at the first position of the loop, that is, ceHP, mm-ceHP, and ceHP-CUUCloop, and thus displayed a similar and almost as strong selectivity as the *C. elegans* RBD. A similar but less pronounced preference for RNAs with a C at the first position of the loop was also found for mut3 (data not shown). In contrast, the mut9 protein showed a strong preference for mmHPs and ce-mmHP over mm-ceHP and ceHP RNA. It also displayed a preference for ceHP-Uloop over ceHP-CUUCloop RNA, but both this preference and the absolute binding to ceHP-Uloop RNA were less pronounced in comparison to ce-mmHP and mmHPs RNAs.

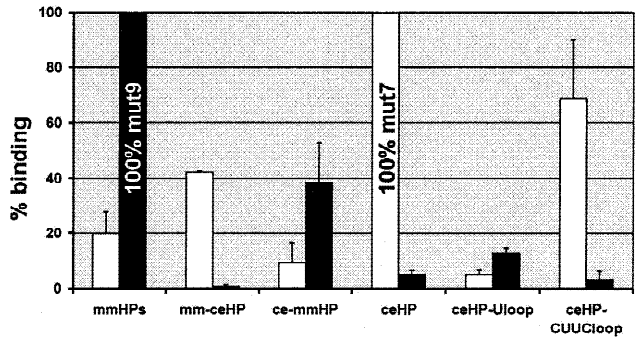

FIGURE 8. Summary of the binding of mut7 (white bars) and mut9 (black bars) RBDs to various hairpin RNAs shown in Figure 2. The RNA–protein complexes were quantitated from three independent experiments by PhosphorImager. The experiment indicates that the new selectivity of both mutant RBDs is mostly for the pyrimidine at the first position in the hairpin loop.

We have therefore identified two sequence elements, YGKNT(24–28) and FKKY(55–58), in the human RBD that are important for RNA–protein interaction. Changing these elements to the respective *C. elegans* sequences increased the ability of the RBD to discriminate between the vertebrate and *C. elegans* hairpin RNAs albeit with opposite effects. The experiments also indicated that both elements affect RNA-binding selectivity mostly through the first nucleotide in the loop.

DISCUSSION

RNA binding of *C. elegans* HBP and its RBD is highly selective for a cytosine at the first position of the histone hairpin loop

So far, HBP (or SLBP) cDNAs have been cloned from three vertebrates (human, mouse, and *Xenopus laevis*) and from the nematode worm *C. elegans* (Wang et al., 1996, 1999; Martin et al., 1997). All these proteins bind very strongly to the 3′ end of their respective histone mRNAs. The RBD, defined for human HBP by deletion analysis (Wang et al., 1996), is the most conserved region among the different proteins (Müller & Schümperli, 1997), but does not show obvious homologies with other known RNA-binding motifs (Burd & Dreyfuss, 1994; Draper, 1995, 1999). A main objective of this work was, therefore, to obtain first insights into what determines the specificity of this RNA–protein interaction.

The binding target for HBP, the hairpin structure at the 3′ end of animal histone mRNAs, has been highly conserved in evolution (Marzluff, 1992; Marzluff & Hanson, 1993; Wittop Koning & Schümperli, 1994; Müller & Schümperli, 1997). However, the sequences at the 3′ end of *C. elegans* histone mRNAs show a number of important deviations from those of other metazoans. First, not only the hairpin and the following 4 nt, but also 12 nt preceding the hairpin are virtually invariant among *C. elegans* histone genes (Wittop Koning & Schümperli, 1994). In comparison, the conserved region preceding the hairpin is shorter and more variable in vertebrates (the main characteristic of this region being the absence of Gs). Second, the sequence flanking the hairpin on the 3′ side is ACA$^{A}/_{U}$ in *C. elegans*, ACCA in sea urchins and ACCCA in vertebrates. Moreover, the *C. elegans* histone hairpins all have a C at the first position of the loop, whereas an invariant U is found at this position in all other metazoans. Finally, *C. elegans* seems to lack a well-conserved downstream or spacer element. The best candidate so far is the sequence $^{A}/_{U}$AAUCC (Wittop Koning & Schümperli, 1994), which has, however not been confirmed experimentally, since the *C. elegans* U7 snRNA, supposed to interact with this sequence, has not yet been identified.

Here we have exploited the availability of cloned cDNAs for human and *C. elegans* HBP to analyze if the two proteins display similar or different RNA binding specificities. Interestingly, the *C. elegans* HBP bound very selectively to its own, but not to a vertebrate hairpin, and this high selectivity was preserved when only the *C. elegans* RBD was analyzed (Fig. 3). This selectivity was not the result of misfolding of the *C. elegans* RBD, as it bound to ceHPs RNA with similar affinity as the human RBD (Fig. 4) and because the binding selectivity was observable over a wide range of salt concentrations (Fig. 7). We further showed that the nucleotide at the first position was mostly responsible for the exclusive binding of the *C. elegans* HBP and its RBD to its cognate RNA hairpin and that this nucleotide had to be a C (Fig. 5). In contrast, exchanging the sequences flanking the hairpin only slightly reduced the binding of the *C. elegans* RBD.

This highly selective binding behavior of the *C. elegans* HBP and its RBD contrasted with that of human HBP. Both human HBP and its RBD bound about equally well to all the mammalian, *C. elegans,* and chimeric hairpins analyzed (Figs. 3 and 5). Previous studies had shown that the sequence of the stem and the sequences immediately flanking the hairpin were critical determinants for mammalian HBP binding but that the conformation of the loop was not critical (Pandey et al., 1991, 1994; Williams & Marzluff, 1995). However, when the conserved uridine at the first position of the loop was mutated to guanosine, the resulting RNA showed only 15% binding to HBP present in mouse myeloma nuclear extract (Williams & Marzluff, 1995). We recently found a similar but slightly weaker reduction in binding of recombinant human HBP to the same hairpin mutation (Martin et al., 2000). Thus mammalian HBPs are apparently sensitive to substitution of the first U in the loop by a purine, whereas a U-to-C substitution is well tolerated (Fig. 5). Moreover, our present data demonstrate that the sequences flanking the vertebrate hairpin can be replaced by their *C. elegans* counterpart without changing the binding behavior of the human RBD. However, it must also be mentioned that other residues, notably at the base of the stem, are very critical for binding of mammalian HBP to the RNA hairpin (Pandey et al., 1994; Williams & Marzluff, 1995; Martin et al., 2000).

The high selectivity of the *C. elegans* HBP and its RBD for the nature of the pyrimidine at the first position of the hairpin loop is rather surprising. Uracil and cytosine differ only in two side groups of the pyrimidine ring and, as discussed above, mammalian HBP binds equally well to hairpin RNAs containing either type of pyrimidine. This strongly suggests that the *C. elegans* RBD must make additional, specific contacts with the cytidine in the first position of the loop and/or form structures that prevent the binding of hairpins containing a different nucleotide at this position.

Two sequence elements in the RBD conferring selectivity for the first position of the histone hairpin loop

Based on a consensus sequence derived from alignment of the five known HBP sequences across the highly conserved RBD (Fig. 1), we converted single amino acids or short oligopeptides from the human to the *C. elegans* RBD sequence by site-directed mutagenesis. This mutational analysis was carried out to identify amino acids that may mediate the highly selective RNA binding displayed by the *C. elegans* RBD. Our results revealed two sequence elements in the human RBD, YGKNT (amino acids 24–28) and FKKY (55–58), that, when changed to the corresponding *C. elegans* sequence, led to a more restrictive binding selectivity (Figs. 6–8).

By converting the sequence YGKNT to RAKEK (mut3 and mut7), we observed a decreased binding to the mammalian RNA compared to the *C. elegans* hairpin. A less pronounced selectivity was obtained in mut13 and mut16 where YGKNT was converted to RAKET and RGKET, respectively, whereas the binding behavior of mut14 (RGKEK) was not changed compared to the original human RBD (Fig. 6). Thus, although several changes in this sequence affected RNA binding selectivity, these effects were variable, depending on which amino acids were altered. In contrast, converting the sequence FKKY into LINF (mut9) dramatically increased the preference of this mutant for the mammalian hairpin, and binding to the *C. elegans* RNA was lost almost completely. This latter finding was remarkable, because we had hereby obtained an RBD with a new binding specificity not displayed by either wild-type human HBP or any of the other HBP polypeptides analyzed so far.

All three altered selectivity mutants, mut3, mut7, and mut9, still bound to the preferred RNA hairpins with affinities comparable to the wild-type hs and ce RBDs (Table 1) and the binding selectivity was conserved over a wide range of salt concentrations (Fig. 7). This indicated that we were dealing with true alterations in binding specificity and not simply with misfolding or a general loss of activity of the proteins.

Interestingly, both types of increased selectivities represented by these mutants discriminated between the different RNAs mostly on the basis of the first nucleotide of the loop (Fig. 8; data not shown). The RBD of mut7 (and mut3) bound more selectively to RNAs that had a C at the first position, whereas mut9 preferentially recognized RNA hairpins with a U at the first position of the loop.

To discuss the possible implications of these findings in more detail, it is important to consider briefly other well-defined RNA–protein interactions (reviewed in Burd & Dreyfuss, 1994; Draper, 1995, 1999). A first class of RNA binding proteins, exemplified by HIV Rev and Tat

proteins or the bacteriophage λ N protein, use arginine-rich α-helices or β-ribbons to interact with accessible parts of the major groove of an RNA A-helix. Most of the contacts are electrostatic or involve hydrogen bonds, but stacking interactions, especially with aromatic side chains, are also frequently seen. Other proteins, for example those containing the widespread ribonucleoprotein or RNA recognition motif (RRM), interact with more exposed parts of an RNA such as loops and bulges. The RRM forms a β-sheet surface that contacts the RNA mostly through positively charged and aromatic side chains of solvent-exposed amino acids (Nagai et al., 1990; references in Burd & Dreyfuss, 1994).

We have previously analyzed HBP using the PHDsec program (Rost & Sander, 1993, 1994) that allows the prediction of protein secondary structure, based on phylogenetic comparisons of primary amino acid sequences. This program proposes with high confidence the presence of three α-helical regions within the RBD, two closely spaced 12 and 9 amino acid long helices in the N-terminal half connected by a loop with a 15 amino acid helix in the C-terminal half (Martin et al., 2000). Interestingly, the two sequence elements identified in this work are localized in the strands connecting the proposed helices (Fig. 6). The N-terminal YGKNT sequence is located between the putative helices 1 and 2, and the C-terminal sequence FKKY is closer to helix 3. In contrast, five of six previously isolated mutations that eliminated or reduced binding of the full-length human HBP to mmHPs RNA were located in the predicted α-helical regions 1, 2, and 3 (Martin et al., 2000). In the context of this secondary structure prediction, it is possible that one or more of the proposed helices contact the major groove of the short RNA stem region and that one or both identified elements in the putative connecting strands interact with the presumably exposed base at the first position of the RNA loop, as is suggested by our present data.

Considering the chemical nature of the two amino acid sequence elements identified in this paper, both the human and the *C. elegans* version of the N-terminal element contain positively charged and aromatic amino acids (YGKNT and RAKEK, respectively; corresponding residues underlined). Thus both versions are good candidates for making specific contacts with RNA. In contrast only the human version of the C-terminal element is rich in positively charged and aromatic amino residues (FKKY), whereas the *C. elegans* sequence contains only one aromatic amino acid (LINF). Thus, changing the sequence FKKY to LINF may have caused a loss of protein–RNA contacts. If this is the case, however, how could this alteration have brought about a new specificity for the mammalian version of the hairpin with U at the first position of the loop?

One possibility is that the human sequence YGKNT recognizes preferentially a U residue at the first position of the loop, whereas the corresponding *C. elegans* element RAKEK binds preferentially to hairpin RNAs with a C at the first position. Consistent with this, mut7 and mut3 bind more strongly to RNAs whose loop begins with C. We further postulate that the second human RBD sequence FKKY binds to other features of the RNA that are common to human and *C. elegans* hairpins, or generally enhances RNA binding by some other means and that the corresponding *C. elegans* sequence LINF has lost this ability. The mammalian FKKY sequence should therefore increase the binding of the corresponding RBD and allow it to bind to RNAs containing the nonspecific pyrimidine at the first position of the loop as is indeed the case for hs RBD as well as for mut7 and mut3. In contrast, in the presence of the *C. elegans*-specific LINF sequence, we observe an exclusive binding behavior, either to hairpins loops beginning with U when the first element is YGKNT (mut9) or to loops beginning with C when the first element is RAKEK (ce RBD).

In support of this concept, the sequence YGKNT is conserved in all four vertebrate HBP sequences, whereas the FKKY sequence is conserved in mouse HBP and *Xenopus* SLBP1, but not in *Xenopus* SLBP2, where the corresponding sequence is SKKY. However, it remains possible or even likely that other amino acids differing between the human and *C. elegans* RBDs also play a role in RNA recognition.

In conclusion, we have identified two amino acid sequence elements that contribute to the sequence selectivity displayed by the RBD of *C. elegans* HBP. Most of the binding selectivity obtained when changing either of these elements from the human to the *C. elegans* sequence is due to discrimination for the pyrimidine present in the first position of the histone hairpin loop. Future experiments aimed at defining which chemical groups of the RBD interact with this important nucleotide may serve as a starting point for a structural model of the HBP-hairpin complex.

MATERIALS AND METHODS

Nucleic acids

The hairpin RNAs (nucleotides of the stem are underlined) mmHPs (5′-GGACAAAAGGCCCUUUUCAGGGCCACCC; previously called wtHPs RNA (Martin et al., 2000)), ceHP (5′-GGAACCGAACCCAACGGCCCUCUUUAGGGCCACAA), mm-ceHP (5′-GGACAAAAGGCCCUCUUUAGGGCCACCC), ce-mmHP (5′-GGAACCGAACCCAACGGCCCUUUUCAGGGCCACAA), ceHP-CUUCloop (5′-GGAACCGAACCCAACGGCCCUCUUCAGGGCCACAA), ceHP-Uloop (5′-GGAACCGAACCCAACGGCCCUUUUUAGGGCCACAA) (Fig. 2B), and ceHPs (5′-GGAACCCAACGGCCCUCUUUAGGGCCACAA) (see Fig. 2 and text) were transcribed from partly double-stranded oligonucleotides by T7 RNA polymerase (Milligan et al., 1987). Uniformly labeled RNA was made by including α-^{32}P-UTP in the transcription reaction and pu-

rified by denaturing polyacrylamide electrophoresis. RNA concentrations were determined by the amount of α-^{32}P-UTP incorporated into the RNA. RNA was stored in H_2O at −20 °C.

Strains

Plasmids were amplified in *E. coli* strain XL1-blue. The single-stranded DNA for the site-directed mutagenesis was made in *E. coli* strain CJ236 *(dut, ung, thi, relA*; pCJ105[Cmr]).

Site-directed mutagenesis

Site-directed mutagenesis was performed as described (Kunkel, 1985). Mutations in the human RBD were introduced using the oligonucleotides hRBD-E9 (5′-TTCATCTG TGCAAAAGTCAGC), hRBD-QKQ (5′-ATCTCCCTCGATC TCCTCATT), hRBD-YG•NT (5′-GTAGGCAATTTTCTCCTT CGCACGGTTG), hRBD-FKKY (5′-TCGACTAAAGTTTATAA GTTTATTAGG), and hRBD-L69 (5′-CACCTTCCATTTTT TGATTTG). Combinations of the different primers were annealed to single-stranded DNA isolated from phage particles produced by the infection of *E. coli* strain CJ236 transformed with pBluescript KS (−) phagemid containing the human RBD cDNA with helper phage M13KO7. Second-strand synthesis was performed using T4 DNA polymerase and the DNA was then introduced into *E. coli* strain XL1-blue. The introduction of mutations was confirmed by DNA sequencing (see below).

Proteins

Primers hbf9 (5′-GCCTCGAGATGTCTACTGTGC) and hbf11 (5′-TGCTCGAGTCATTCCGCTGGAGGA) and primers hbf9 and hbf12 (5′-TGCTCGAGCAGTTAGCTCATGG) were used to amplify from human HBP cDNA (Martin et al., 1997) the RBD (hs RBD) and the RBD plus the C-terminal region (hs RBD-Ct), respectively (Fig. 3A). Primers celeg-1 (5′-CACGA AGACTCGAGATGGAAGAGCCGAC) and celeg-2 (5′-GTAT GGACCTCGAGTCATGGCTCCTCTCC), primers celeg5′ (5′-GGCGCCTCGAGCGCCTAAAAAAATGGC) and celeg-2, and primers celeg-1 and celeg3′ (5′-GAAGGAACTCGAGTAT CTTTAGTGCGACG) were used to amplify from *C. elegans* HBP cDNA (Martin et al., 1997) the corresponding ce RBD, ce Nt-RBD, and ce RBD-Ct fragments, respectively. Amplification products were cut with *Xho*I, gel-purified, and ligated into the *Xho*I site of pBluescript KS (−). The template for hs Nt-RBD had previously been isolated as a nonsense mutation at Q208 in the human HBP cDNA (Martin et al., 2000).

Proteins were synthesized in vitro in wheat germ extract as described by the supplier (Promega) in 50 μL, using 1.5 μg of linearized plasmids. For analysis by SDS-PAGE, ^{35}S-methionine-labeled proteins were precipitated by trichloroacetic acid (10–15% v/v), washed twice with acetone, and resuspended in 20 μL of SDS-PAGE loading buffer. Proteins were analyzed by 15% SDS-PAGE. Detection of proteins was by PhosphorImager.

High-yield protein synthesis for the experiments described in Figures 4 and 7 and Table 1 was done with RNA prepared in vitro using T7 RNA polymerase. For translation in wheat germ extract, the protocol of the manufacturer (Promega) was modified such that the concentrations of the amino acids was 120 μM, except for leucine, which was at 80 μM, and methionine, which was at 40 μM. Synthesis was done in 100 μL in the presence of ^{35}S-methionine for 2 h as described by the manufacturer. Proteins were analyzed by 12.5% Tris-tricine gel electrophoresis (Schägger & von Jagow, 1987) and visualized by autoradiography of dried gels. Protein bands were excised and the fraction of ^{35}S methionine incorporated into protein was determined by comparison with an aliquot of the wheat germ reaction mixture. Protein yields were calculated taking into account the numbers of methionine present in each product and the protein concentrations in the translation mixture were 18.4, 20.5, 16.4, 22, and 25.3 nM for hs RBD, mut3, mut7, mut9, and ce RBD, respectively.

Sequencing

DNA sequencing was done with the oligonucleotide T7 primer (5′-AATACGACTCACTATAG) labeled at the 5′-end with IRD41, using the Thermo Sequenase fluorescent labeled primer cycle sequencing kit with 7-deaza-dGTP (Amersham). Sequences were analyzed on an automatic sequencer (Licor 4200).

Binding assays

Twenty-seven femtomoles ^{32}P-labeled RNA were incubated with 5 μL of the in vitro-made protein in a final volume of 10 μL of 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM DTT, 10% glycerol, 1 μg/μL yeast tRNA, 1 unit/μL RNasin (Promega) for 20 min on ice. As a control, RNAs were incubated in the same buffer with 5 μL of translation mix that had been incubated without plasmid. The reaction products were analyzed directly by EMSA as described (Schaller et al., 1997). Products were visualized by autoradiography or by PhosphorImager (Molecular Dynamics).

For the determination of affinity constants in Figure 4 and Table 1, 5 μL binding reactions containing tRNA and RNasin as described above were set up using 4.4 nM protein prepared by the high-yield synthesis protocol. Where necessary, mock translation mixture was added to keep the amount of wheat-germ extract added to the reaction constant. Either mmHPs or ceHPs RNA was then added at the indicated concentration and the mixture was incubated on ice and subsequently analyzed by EMSA as described above, except that a Fuji PhosphorImager was used for quantitation.

For the salt titration shown in Figure 7, reactions were set up at 50 mM KCl or without KCl (~15 mM K acetate is carried over from the translation mixture) using 4.4 nM protein prepared by the high-yield synthesis protocol and 4.9 nM RNA. Subsequently, 3.5-μL aliquots were either left unchanged or supplemented with KCl to 100, 200, 400, 600, or 1,000 mM, and incubated for a further 20 min on ice prior to analysis by EMSA as described above.

ACKNOWLEDGMENTS

This work was supported by the state of Bern, Swiss National Science Foundation Grant 31-52619.97 and the University of Aberdeen. We thank Franck Martin for important advice throughout this project and Oliver Mühlemann for critical comments on the manuscript.

Received March 7, 2000; returned for revision April 6, 2000; revised manuscript received July 25, 2000

REFERENCES

Bond UM, Yario TA, Steitz JA. 1991. Multiple processing-defective mutations in a mammalian histone pre-messenger RNA are suppressed by compensatory changes in U7 RNA both in vivo and in vitro. *Genes & Dev 5*:1709–1722.

Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science 265*:615–621.

Cotten M, Gick O, Vasserot A, Schaffner G, Birnstiel ML. 1988. Specific contacts between mammalian U7 snRNA and histone precursor RNA are indispensable for the in vitro 3′ RNA processing reaction. *EMBO J 7*:801–808.

Dominski Z, Sumerel J, Hanson RJ, Marzluff WF. 1995. The polyribosomal protein bound to the 3′ end of histone mRNA can function in histone pre-mRNA processing. *RNA 1*:915–923.

Draper DE. 1995. Protein–RNA recognition. *Annu Rev Biochem 64*:593–620.

Draper DE. 1999. Themes in RNA–protein recognition. *J Mol Biol 293*:255–270.

Gick O, Krämer A, Vasserot A, Birnstiel ML. 1987. Heat-labile regulatory factor is required for 3′ processing of histone precursor mRNAs. *Proc Natl Acad Sci USA 84*:8937–8940.

Kunkel TA. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci USA 82*:488–492.

Martin F, Michel F, Zenklusen D, Müller B, Schümperli D. 2000. Positive and negative mutant selection in the human histone hairpin-binding protein using the yeast three-hybrid system. *Nucleic Acids Res 28*:1594–1603.

Martin F, Schaller A, Eglite S, Schümperli D, Müller B. 1997. The gene for histone RNA hairpin binding protein is located on human chromosome 4 and encodes a novel type of RNA binding protein. *EMBO J 15*:769–778.

Marzluff WF. 1992. Histone 3′ends: Essential and regulatory functions. *Gene Expr 2*:93–97.

Marzluff WF, Hanson RJ. 1993. Degradation of a nonpolyadenylated messenger: Histone mRNA decay. In: *Control of messenger RNA stability.* New York: Academic Press. pp 267–290.

Meier VS, Böhni R, Schümperli D. 1989. Nucleotide sequence of two mouse histone H4 genes. *Nucleic Acids Res 17*:795.

Melin L, Soldati D, Mital R, Streit A, Schümperli D. 1992. Biochemical demonstration of complex formation of histone pre-mRNA with U7 small nuclear ribonucleoprotein and hairpin binding factors. *EMBO J 11*:691–697.

Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res 15*:8783–8798.

Mowry KL, Steitz JA. 1987a. Identification of the human U7 snRNP as one of several factors involved in the 3′ end maturation of histone pre-messenger RNAs. *Science 238*:1682–1687.

Mowry KL, Steitz JA. 1987b. Both conserved signals on mammalian histone pre-mRNAs associate with small nuclear ribonucleoproteins during 3′ end formation in vitro. *Mol Cell Biol 7*:1663–1672.

Müller B, Schümperli D. 1997. The U7 snRNP and the hairpin binding protein: Key players in histone mRNA metabolism. *Semin Cell Dev Biol 8*:567–576.

Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature 348*:515–520.

Pandey NB, Sun JH, Marzluff WF. 1991. Different complexes are formed on the 3′ end of histone mRNA with nuclear and polysomal proteins. *Nucleic Acids Res 19*:5653–5659.

Pandey NB, Williams AS, Sun JH, Brown VD, Bond U, Marzluff WF. 1994. Point mutations in the stem-loop at the 3′ end of mouse histone mRNA reduce expression by reducing the efficiency of 3′ end formation. *Mol Cell Biol 14*:1709–1720.

Roberts SB, Emmons SW, Childs G. 1989. Nucleotide sequences of *Caenorhabditis elegans* core histone genes. *J Mol Biol 206*:567–577.

Rost B, Sander C. 1993. Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol 232*:584–599.

Rost B, Sander C. 1994. Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins 19*:55–72.

Schägger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem 166*:368–379.

Schaller A, Martin F, Muller B. 1997. Characterization of the calf thymus hairpin-binding factor involved in histone pre-mRNA 3′ end processing. *J Biol Chem 272*:10435–10441.

Soldati D, Schümperli D. 1988. Structural and functional characterization of mouse U7 small nuclear RNA active in 3′ processing of histone pre-mRNA. *Mol Cell Biol 8*:1518–1524.

Vasserot AP, Schaufele FJ, Birnstiel ML. 1989. Conserved terminal hairpin sequences of histone mRNA precursors are not involved in duplex formation with the U7 RNA but act as a target site for a distinct processing factor. *Proc Natl Acad Sci USA 86*:4345–4349.

Wang ZF, Ingledue TC, Dominski Z, Sanchez R, Marzluff WF. 1999. Two *Xenopus* proteins that bind the 3′ end of histone mRNA: Implications for translational control of histone synthesis during oogenesis. *Mol Cell Biol 19*:835–845.

Wang ZF, Whitfield ML, Ingledue TC, Dominski Z, Marzluff WF. 1996. The protein that binds the 3′ end of histone mRNA: A novel RNA-binding protein required for histone pre-mRNA processing. *Genes & Dev 10*:3028–3040.

Williams AS, Marzluff WF. 1995. The sequence of the stem and flanking sequences at the 3′ end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. *Nucleic Acids Res 23*:654–662.

Wittop Koning TH, Schümperli D. 1994. RNAs and ribonucleoproteins in recognition and catalysis. *Eur J Biochem 219*:25–42.