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

Anthony S. Williams and William F. Marzluff*

Program in Molecular Biology and Biotechnology, Department of Biochemistry and Biophysics, University of North Carolina 27599, Chapel Hill, NC 27599, USA

Received October 5, 1994; Revised and Accepted January 4, 1995

ABSTRACT

Complexes of different electrophoretic mobility containing the stem-loop binding protein, a 45 kDa protein, bound to the stem-loop at the 3' end of histone mRNA, are present in both nuclear and cytoplasmic extracts from mammalian cells. We have determined the effect of changes in the loop, in the stem and in the flanking sequences on the affinity of the SLBP for the 3' end of histone mRNA. The sequence of the stem is particularly critical for SLBP binding. Specific sequences both 5' and 3' of the stem-loop are also required for high-affinity binding. Expanding the four base loop by one or two uridines reduced but did not abolish SLBP binding. RNA footprinting experiments show that the flanking sequences on both sides of the stem-loop are critical for efficient binding, but that cleavages in the loop do not abolish binding. Thus all three regions of the RNA sequence contribute to SLBP binding, suggesting that the 26 nt at the 3' end of histone mRNA forms a defined tertiary structure recognized by the SLBP.

INTRODUCTION

Histone mRNAs are the only mRNAs in metazoans which lack a polyA tail, ending instead with a stem-loop structure which has been highly conserved (1,2). Since replication-dependent histone genes do not contain introns, the only processing step in histone mRNA biosynthesis is a 3' end endonucleolytic cleavage after the stem-loop (3). In mammals this processing reaction requires three factors: the U7 snRNP which binds to a 'purine rich' element 10–20 nt downstream of the stem-loop sequence (4–6); a hairpin binding factor (7–9), which recognizes the stem-loop, and a heat-labile factor whose function is unknown (10). The stem-loop is not absolutely essential for processing *in vitro*, but the requirement for the stem-loop in some genes can be replaced by increasing the number of base pairs between the histone pre-mRNA and U7 snRNP (11). Following processing the histone mRNA is exported to the cytoplasm and there is evidence that the histone 3' end is required for efficient transport (12,13) and translation (14). The histone 3' end is also necessary and

sufficient for regulation of the half-life of histone mRNA (15). The levels of replication-dependent histone mRNAs are regulated during the cell cycle at both the transcriptional and post-transcriptional level (16–18). Both the processing of the histone pre-mRNA (19,20) and the half-life of the cytoplasmic polyribosomal histone mRNA (17) are regulated and the 3' end of the histone mRNA is critical to both regulatory events. The post-transcriptional regulation, directed by the histone 3' end, accounts for the majority of the regulation of histone mRNA levels (18–21). The histone 3' end is unusual among RNA structures in that not only have the single-stranded loop regions of the RNA been conserved, but the sequence of the stem as well as the sequences flanking the stem have also been conserved (2). Using a mobility shift assay and UV-crosslinking, we have detected a 45 kDa protein, termed the stem-loop binding protein (SLBP), which is present in both the nucleus and polyribosomes of mammalian cells (22). It is possible that the nuclear SLBP complex is part of the hairpin binding factor required for histone pre-mRNA processing (23). Here we report that all three of the sequence elements, the stem, the loop and the flanking sequences, are required for high affinity binding of the SLBP.

MATERIALS AND METHODS

Gel mobility shift assays

The gel mobility shift assay was performed as described previously (22). The various mutant stem-loop RNA probes were synthesized by T7 RNA polymerase and [α - 32 P]UTP using the appropriate oligonucleotide template (24). In some experiments unlabelled RNA was synthesized, treated with calf-intestinal phosphatase and then labeled at the 5' end with [γ - 32 P]ATP and polynucleotide kinase. All RNAs were purified by preparative gel electrophoresis. $3\text{--}4 \times 10^{-15}$ moles of wt or mutant stem-loop RNA probe was incubated with 12 μ g of mouse myeloma nuclear proteins or 2 μ g of a 0.8 M KCl extract of mouse myeloma polyribosomal proteins for 20 min on ice as previously described (22). The complexes were resolved by electrophoresis on a 6% non-denaturing gel containing 3.3 mM NaOAc, 1 mM EDTA and 6.7 mM Tris-HCl (pH 7.5) and quantified on a Phosphor-Imager (Molecular Dynamics).

* To whom correspondence should be addressed

For some experiments RNA (either with ribose or 2'-OCH₃ ribose) was synthesized on an oligonucleotide synthesizer and then labeled with [γ -³²P]ATP and polynucleotide kinase. RNA was also synthesized with biotin on the 3' end using a CPG column containing biotin (Glen Research). All oligonucleotides (DNA and RNA) were synthesized in the Oligonucleotide Facility of the Lineberger Comprehensive Cancer Center at the University of North Carolina.

Competition assays

For competition experiments, between 3 and 4 $\times 10^{-15}$ moles of wt stem-loop RNA probe was used in each binding reaction. The probe and the competitor RNAs were mixed prior to addition of the extract. 10 μ g of nuclear proteins from exponentially growing mouse myeloma cells (4–6 $\times 10^5$ cells/ml) were incubated with the wt RNA stem-loop probe and varying amounts of the unlabeled competitors for 20 min and then analyzed by gel electrophoresis. In each set of experiments a competition curve with wild-type RNA was done in parallel with the mutant RNA. The relative affinities of the mutant RNA for the SLBP was estimated by the amount of competitor required to give 50% competition of complex formation.

A variation of the competition assay was used to determine the half-life of the SLBP-RNA complex once equilibrium has been reached. γ -³²PO₄-labeled SL-wt RNA was incubated with 12 μ g of total nuclear proteins for 20 min. Then a 1000-fold excess of unlabeled SL-wt RNA was added to each reaction and at various times between 1 min and 6 h, each reaction was loaded on a 6% nondenaturing polyacrylamide gel. The amount of the SLBP-RNA complex and unbound RNA was determined using a PhosphorImager (Molecular Dynamics).

RNA footprinting

A 5' labeled probe was made from wild-type unlabeled RNA as described above. The 3' labeled probe was made by incubating the wt RNA with [³²PO₄]pCp and T4 RNA ligase. The labeled RNAs were purified by gel electrophoresis. The probe was treated with 75 mM KOH for 8 min at 42°C in order to generate a cleavage ladder and the solution was neutralized with an equivalent amount of glacial acetic acid. 3–4 $\times 10^{-15}$ mol of the RNA ladder was incubated for 20 min with 12 μ g of nuclear proteins from exponentially growing mouse myeloma cells. Gel mobility shift assays were performed using a 6% non-denaturing gel. After autoradiography, the RNA-SLBP complex was cut out and the RNA recovered by electroelution. The RNAs were precipitated with EtOH, and analyzed by electrophoresis on a 20% polyacrylamide–7 M urea gel together with the appropriate RNA ladder. For RNase U2 digestion the 5' end-labeled RNA was mixed with 1.5 μ g of yeast tRNA in 1.75 M urea, 8 mM sodium citrate (pH 3.5), 0.4 mM EDTA and treated with 0.5 U of U2 RNase at 55°C for 12 min. For RNase A digestion, the labeled RNA was mixed with 10 μ g of yeast tRNA in 1.75 M urea, 10 mM Tris-HCl and treated with 0.001 U of RNase A for 2 min at 25°C. The samples were chilled on ice and immediately loaded on the gel.

RESULTS

The stem-loop sequence has been highly conserved in all metazoan histone genes (2). The consensus sequence of the 3' end of vertebrate histone mRNAs, derived from over 100 histone

Figure 1. Wild-type (wt) and mutant 3' end histone RNA stem-loop structures. (A) The consensus sequence of the 3' end of histone mRNA taken from 100 vertebrate histone genes is shown. The boxed nucleotides in the loop and base pairs in the stem are invariant. (B) The mutant stem-loop structures are shown with the mutations in lower case. The SL(RS) and SL(CG/CG) have the entire stem or the bottom two conserved GC/GC base pairs reversed respectively and the altered nucleotides in these structures are boxed or circled. SL(U₄C) and SL(U₅C) mutant structures have one extra uracil and two extra uracils in the loop respectively.

gene sequences from mouse, rat, chicken, human and frog, is shown in Figure 1A. A similar consensus is found if the comparison is extended to all metazoans, a total of 400 sequences from diverse phyla, including annelids, echinoderms, and insects. The highly conserved nucleotides include the entire 16 nt stem-loop and the 5 nt 5' and 3' of the stem-loop. In vertebrates there are invariant uridines in the first and third position of the loop. The third position is invariant in all metazoans, while the first position is a uridine in all metazoans, except *C.elegans*, which has a cytosine in the first position (25). The second position is predominantly a uridine although cytosines and adenosines are found in this position in some genes. The fourth position is either a U, C or A. There is not a G in the loop in any histone gene sequence reported. The 6 bp stem has an invariant UA base pair at the top of the stem and two invariant GC base pairs at the base

of the stem. The three base pairs between these conserved base pairs always have pyrimidines on the 5' side of the stem, and there are at least two C-G base pairs.

The 5 nt flanking either side of the stem-loop of all vertebrate histone mRNAs have been as conserved as the stem-loop sequence (2). The 5' flanking nucleotide consensus is CCAA, with the consensus nucleotide at each position present over 80% of the time. More than 95% of the nucleotides at all five positions are C's or A's. The 3' flanking consensus is ACCCA and again >95% of the nucleotides are C's or A's. While occasional uridines are found, there are almost never G's in these 10 flanking nucleotides (<1%). The 3' end of histone mRNA is formed by a cleavage reaction after a CA 3-6 nt from the end of the base of the stem-loop (11,26).

Previously we have focused on the effect of point mutations in the loop on the binding of the SLBP and the expression of the histone genes (23), measuring the binding of the SLBP by the formation of a specific complex in a mobility shift assay. Here we have varied all parts of the consensus sequence to better assess the requirements for SLBP binding. The sequence of the mutant probes are shown in Figure 1B. In addition we tested the binding to DNA with the same sequence and to methylated RNA, containing 2'OCH₃ ribose at each position.

A sensitive assay to determine the relative affinities of mutant stem-loop sequences for the SLBP is to measure the ability of the mutant RNA to compete for binding of the wild-type stem-loop sequence. In this assay ³²P-labeled SL_{wt} RNA was mixed with varying amounts of unlabeled wt stem-loop RNA or mutant stem-loop RNAs prior to analysis of the amount of complex formed by electrophoresis. The relative affinities of the RNAs for the mutant sequence is determined by comparing the amount of mutant competitor and wild-type competitor required to reduce the amount of SLBP complexes by 50%. Thus mutants SL(G¹) and SL(A³) showed 5-10-fold reduced binding as assayed by formation of a complex by mobility shift (22) and a similar 5-8-fold reduction in affinity in the competition assay (data not shown). Thus combining the direct mobility shift assay and the competition assay we can assay relative affinities of altered structures for the SLBP over several orders of magnitude. Note that these binding assays are done in the presence of a large amount of non-specific competitor tRNA, and the addition of even large amounts of specific competitors did not significantly affect the total RNA concentration in the binding assay. For mutant RNA structures which can form a complex detected by the mobility shift assay, the competition assay gives results similar to the mobility shift assay.

An example of the application of this assay to determine relative affinities of different RNAs for the SLBP is shown in Figure 2. Of the mutants shown in Figure 2, the SL(A^{1,3}) and the SL(RS) mutants did not give detectable binding in the mobility shift assay, while the SL(CG/CG) mutant gave less than 5% of the wild-type binding as judged by the mobility shift assay (22,23).

Mutating the conserved bases in the loop and reversing base pairs in the stem reduce SLBP binding to different extents

Previously we have shown, using a mobility shift assay, that changing the conserved first and the third bases in the loop from uridines to adenosines [SL(A^{1,3})] resulted in no detectable SLBP-RNA complex formation with either nuclear or polyribosomal proteins as judged by a gel-mobility shift assay (22). The

Figure 2. Relative affinity of the wild-type (wt) and mutant stem-loop RNA structures for the SLBP. The wt RNA probe was made by T7 RNA polymerase and an oligonucleotide template as described previously (22). The unlabeled wt or mutant competitor RNAs were made in the same manner except the oligonucleotide sequences were changed to produce the mutant RNAs. The synthetic wt RNA probe was mixed with increasing amounts of competitor (1-10 000-fold molar excess) before addition of 12 µg of nuclear proteins from mouse myeloma cells and 60 µg of yeast tRNA. The complexes were resolved by polyacrylamide gel electrophoresis under nondenaturing conditions and detected by autoradiography. The sequence of the competitor RNA is shown below each panel. Lane 1 in each case is the complex formed in the absence of competitor. (A) Competition by the wild-type stem loop (lanes 2-6). (B) Competition by the SL(A^{1,3}) RNA (lanes 2-6). (C) Competition by the SL(RS) RNA (lanes 2-6). (D) Competition by the SL(CG/CG) RNA (lanes 2-6). The band labeled N represents the specific nuclear complex, and band labeled ns represents a non-specific complex. The band labeled RNA is the free probe.

same result was obtained with the SL(RS) which has the base pairs of the entire stem reversed (23). The ability to detect a complex by the gel mobility shift assay is limited by the affinity

Figure 3. Relative affinity of mutant stem-loops for the SLBP. The nuclear extract was incubated with the wild-type probe in the presence of various amounts of competitor RNAs and analyzed as in Figure 2. The amount of wild-type probe bound with no added competitor was set at 100%. The amount of specific complex formed at each concentration was determined using a PhosphorImager. The competitors used are shown in the inset.

of the SLBP for the RNA, which must be high enough to survive during electrophoresis. Hence it is not possible to judge the relative binding affinities of sequences which do have high enough affinities to form a detectable complex. In the competition assay we compare the amounts of the unlabeled mutant and wild-type RNA necessary for displacement of the wild-type probe from the complex and this ratio is a reflection of the relative binding affinities of the two sequences. We compared the affinities of the two sequences, [SL(A^{1,3}) and SL(RS)], which did not form detectable complexes using the mobility shift assay. The unlabeled wild-type sequence competed efficiently as predicted: a 10-fold molar excess reduced binding (Figs 2A and 3) and higher concentrations effectively prevented binding to the labeled probe. The SL(A^{1,3}) RNA also competed for binding to the labeled probe but only at much higher concentrations; binding to the wild-type probe was only effectively competed with a 1000-fold molar excess of unlabeled SL(A^{1,3}) RNA (Fig. 2B). Since 20–40 times as much SL(A^{1,3}) RNA as wild-type RNA was necessary to achieve a similar degree of competition, we conclude that the affinity of the SL(A^{1,3}) RNA for the SLBP is 2.5–5% of the affinity of the wild-type RNA, even though the SL(A^{1,3}) RNA does not form a complex stable enough to be detected in a mobility shift assay.

In contrast the SL(RS) RNA competed much less efficiently for binding with only 50% competition at a 10 000-fold molar excess (Fig. 2C, lane 6; Fig. 3). Thus the SL(RS) RNA, which has a wild-type loop and flanking sequence, has a much lower affinity (<10⁻³) for the SLBP than the wild-type sequence and about a 50-fold lower affinity than the SL(A^{1,3}) RNA. Changing the nucleotides in the stem while maintaining all the base pairs had a much larger effect than altering the two invariant nucleotides in the loop.

To further characterize the requirements for the stem sequence, the two conserved GC base pairs at the base of the stem were reversed [mutant SL(CG/CG), Fig. 1B], and the relative affinity of this mutant RNA for the SLBP was determined. There was less than 5% of the complex formed on the SL(CG/CG) RNA compared to the wild-type as measured by a mobility shift assay [(23) and data not shown]. About 20–50 times more SL(CG/CG) competitor than wild-type RNA was needed to compete the binding of the SLBP to the wild type stem-loop (Fig. 2D). Thus inverting just two base pairs at the base of the stem had a dramatic

effect on SLBP binding, similar to changing the two invariant nucleotides in the loop. The [SL(CG/CG)] mutation also abolished expression of histone mRNA *in vivo* (23).

Other mutations in the stem-loop which affect SLBP binding have been described previously. Mutations changing the conserved base pair at the top of the stem from UA to CG or UG had much smaller effects than reversing the stem sequence, reducing binding 5-fold (CG) to 10–20-fold (UG), although a detectable complex was still formed with both probes (22,23). Competition experiments confirmed the 5–10-fold reduction in binding (data not shown). The effects of a large number of mutations in the stem-loop sequence on binding of the SLBP are summarized in Table 1. A graph of representative competition data is shown in Figure 3.

Table 1. Relative affinity of the mutant stem-loops for the SLBP

Loop mutations		Stem mutations		Flanking mutations	
A ³	20	CG	20	SL(Δ3')	5
G ¹	15	UG	7	SL(Δ5')	3
A ^{1,3}	5	RS	0.3	SL(Δ5'Δ3')	1
U ₄ C	20	UU	<5	5'UUUG	<5
U ₅ C	10	CG/CG	3	3'GGGCA	<5
2'OCH ₃	<0.01	UC	<5	(GC) ₃	<5

The affinity of the SLBP for the various stem-loop sequences was estimated by either a gel mobility shift assay or by the ability of the mutant RNA sequence to compete for the binding of the wild-type sequence. Data for some of the mutants is taken from ref. 22. The mutant 2'OCH₃ has all the nucleotides as 2'OCH₃ ribose. The binding affinity for the wild-type sequence was arbitrarily set at 100.

No detectable complex formation was observed with either a 30 nt DNA with the same sequence or with an RNA containing all 2'OCH₃ ribonucleotides. No detectable competition for binding to the wild-type sequence was observed with a 10 000 fold molar excess of either the DNA (not shown) or 2'OCH₃ ribonucleotide (Fig. 3). This result suggests that the 2'OCH₃ RNA assumes a very different structure from the wild-type RNA or that the 2'OH groups are critical for the protein–RNA contacts.

The 5' and 3' flanking sequences are required for efficient SLBP binding

Mutations were made that deleted the 3' flanking region, part of the 5' flanking region or both the 5' and 3' flanking regions [SL(Δ5'), SL(Δ3'), SL(Δ5'Δ3'), Fig. 1] as well as mutations which altered these sequences by including guanines and/or uridines in either the 5' or 3' flanking sequence [SL(5'UUUG) and SL(3'GGGCA); see Fig. 1B].

The SL(Δ5'Δ3') RNA has a GAA sequence 5' of the stem-loop (necessary for synthesis of this RNA with T7 RNA polymerase), and contains two of the conserved 5' flanking nucleotides, and no 3' flanking nucleotides. The SL(Δ5', Δ3') RNA does contain the wild-type stem-loop sequence, but it did not give a detectable complex when analyzed by the mobility shift assay (not shown). When assayed by the competition assay, this RNA had a 50–100-fold lower affinity for the SLBP than the wild-type stem-loop probe (Fig. 4A), an affinity similar to the SL(A^{1,3}) RNA. Thus either the 5' and/or 3' flanking sequences are necessary for high affinity binding. To determine which of the 5' and 3' flanking regions were most important for SLBP binding,

the mutant stem-loop RNAs SL($\Delta 5'$) and SL($\Delta 3'$) were incubated with total nuclear proteins or polyribosomal proteins and a gel mobility assay performed. Figure 4B shows that both flanking regions are required for efficient SLBP binding with either the nuclear or polyribosomal forms of the SLBP. Based on the relative intensities of the mobility shift, the nuclear SLBP (Fig. 4B, compare lane 1 with lanes 2 and 3) and polyribosomal SLBP (Fig. 4B, compare lane 4 with lanes 5 and 6) bound 5–10% as well to the SL($\Delta 3'$) and less than 5% as well to the SL($\Delta 5'$) RNA compared with the wild-type stem-loop RNA. Note that the SL($\Delta 5'$) RNA has all the 3' flanking sequence and a GAA 5' flanking sequence, with the two A's present in the conserved structure. Thus 17 of the 20 nucleotides in the consensus sequence are present, and yet binding is reduced much more than removing the entire 3' flanking sequence. Binding of the SL($\Delta 5'$) could be reduced because of the substitution of the G in the conserved region, or because the remainder of the conserved 5' flanking sequence has been deleted. Thus the 5' flanking region is critical for efficient binding, but both the 5' and 3' flanking regions are necessary for the high affinity SLBP binding.

Specific 5' and 3' flanking nucleotides are required for SLBP binding

It is possible that the requirement for flanking sequences did not reflect a sequence-specific requirement but a general requirement for additional nucleotides flanking the stem-loop. To address this possibility we constructed two mutant RNAs which have altered 5' or 3' flanking sequences. The first mutant, SL(5'AUUUG), contains UUUG instead of the sequence AAAA immediately before the stem loop, and the second mutant, SL(3'GGGCA), contains the 3' sequence GGGCA in place of the wt sequence ACCCA. These sequences were chosen to avoid introducing any other obvious secondary structural features into the stem-loop.

These mutations reduced binding at least as much as the SL($\Delta 3'$) or SL($\Delta 5'$) mutations, as judged by the mobility shift assay with either nuclear or polyribosomal extracts (Fig. 5). There were complexes formed on the SL(5'AUUUG) RNA with the nuclear proteins, but these complexes had a different mobility from the SLBP complexes and, more importantly, were not competed with an excess of wild-type RNA (Fig. 5, lanes 1–3). There was no detectable complex formation with the SL(5'UUUG) probe in the polyribosomal extract (Fig. 5, lane 7). Thus changing the four nucleotides prior to the stem-loop reduced binding more than 20-fold. Both the nuclear and polyribosomal SLBP bound less than 5% as well to the SL(3'GGGCA) RNA as to the wild-type RNA (Fig. 5, lanes 4, 5, 6 and 8). These results show that specific 5' and 3' flanking sequences are necessary for high affinity binding of the SLBP. Both the deletions and mutations of the flanking sequences show a larger effect of altering the 5' flanking sequences compared to altering the 3' flanking sequences.

A 6 bp stem is essential but the loop size may vary

Since the stem contains 6 bp in all metazoan histone genes, a mutant stem-loop RNA [SL(GC)₃, Fig. 1B] containing an extra GC base pair was tested for its effect on binding. Less than 5% as much complex was formed compared to the wt stem-loop RNA with both the nuclear and polyribosomal extracts (Fig. 6, lanes 1, 2, 5 and 6). The SL(GC)₃ RNA has exactly the wild-type

Figure 4. Role of the flanking sequences in SLBP binding. (A) Competition experiments using the wild-type stem-loop RNA probe and unlabeled wild-type (SL_{wt}, lanes 3–7) or the mutant [SL($\Delta 5'$, $\Delta 3'$)] RNA (lanes 9–13) were performed exactly as described in Figure 2. Lane 1 is probe without added extract, and lanes 2 and 8 are incubations without added competitor. (B) Gel mobility shift assays were done using the wild-type or deletions of the 5' or 3' flanking regions of the stem-loop [SL($\Delta 5'$) and SL($\Delta 3'$)] as probes. The mobility shift assay was done using 12 μ g of nuclear proteins (lanes 1–3) or 2 μ g of proteins extracted from polyribosomes (lanes 4–6). The band labeled N represents the specific nuclear complex and the band labeled P represents the specific polysomal complex and the band labeled RNA represents unbound probe. The stem-loop RNA sequences used are shown at the bottom of the gel.

stem-loop structure with a single extra base pair at the base of the stem. This result is convincing evidence for the requirement of precise orientation of the flanking sequences with respect to the stem for efficient binding of the SLBP.

There are always four bases in the loop in metazoan histone genes. We expanded the loop by adding, one [SL(U₄C)] or two [SL(U₅C)] additional uridines creating a 5 or 6 nt loop. Complexes were readily detected with each mutant, indicating that binding was not drastically reduced. The nuclear SLBP bound at least 20% as well to the SL(U₄C) and 10% as well to SL(U₅C) mutant as to the wild-type stem-loop (Fig. 6, lanes 3 and 4). Similar results were obtained with the extract from polyribo-

Figure 5. Requirements for specific 5' and 3' flanking nucleotides for SLBP binding. Gel mobility shift assays were performed with either nuclear (lanes 1–5) or polyribosomal proteins (lanes 6–8) using the wt and mutant stem-loop RNAs which have altered 5' or 3' flanking sequences. The probes used [wild-type, SL(5'AUUUG) or SL(3'GGGCA)] are indicated above each lane. In lane 3 a 1000-fold unlabeled synthetic wt RNA was included in the reaction. The band labeled N represents the specific nuclear complex, the band labeled P represents the specific polysomal complex, and the band labeled RNA represents free probe. The sequences of the probes are given below. The complexes formed in lane 2 are not due to the SLBP since they are not competed by excess wild-type competitor.

some (Fig. 6, lanes 7 and 8). Thus while adding additional uridines resulted in a stepwise reduction in protein binding, the effect was not as dramatic as changing the sequence of the four-base loop or making alterations in the stem or in the flanking sequences. This result suggests that the precise conformation of the loop is not as important as the structure of the stem and is not essential for binding of the SLBP.

The minimal SLBP binding site

The results described above implicate all portions of the stem-loop sequence in high affinity binding to the SLBP. To determine a minimum RNA binding site, we labeled either the 5' end or the 3' end of the wild-type stem-loop RNA and then generated a collection of labeled RNAs differing in size by one nt by partial alkaline hydrolysis. The starting material contains two RNA species of 30 and 31 nt, arising from the addition of a non-templated A to many of the T7 transcripts. This collection of RNAs was incubated with a nuclear extract and the bound RNAs isolated after separation of the SLBP-RNA complex by gel electrophoresis. The population of bound RNAs was then analyzed by polyacrylamide gel electrophoresis. Only those labeled RNAs which bound efficiently to the SLBP will be present in the complex. As shown using the 5' labeled probe, only 2 nt can be removed from the 3' end of the probe without affecting the binding of the SLBP (Fig. 7, lane 4). Surprisingly, although removal of more than 2 nt from the 3' end greatly reduced binding, much of the RNA with cleavages after nucleotides 15,

Figure 6. Effect of the length of the stem and the size of the loop on SLBP binding. The wild-type or mutant [SL(U₄C), SL(U₅C) or SL(GC)₃] probes were incubated with 12 µg of nuclear protein (lanes 1–4) or 2 µg of polysomal protein (lanes 5–8) from mouse myeloma cells in the presence of 60 µg of tRNA and the complexes resolved by polyacrylamide gel electrophoresis under nondenaturing conditions. The band labeled N represents the specific nuclear complex, and the band labeled P represents the specific polysomal complex. The band labeled RNA represents the free probe.

16, 17 or 18 was found in the bound fraction. These cleavage sites map to the loop, as demonstrated using U2 RNase cleavage at A's and RNase A which preferentially cleaved in the loop to precisely identify the bases in the random cleavage ladder (Fig. 7, lanes 5 and 7). The cleavage sites are shown as arrows below the figure. The intensity of the bands due to cleavage in the loop in the bound RNA suggests that these cleaved molecules bound the SLBP nearly as well as the intact RNA, suggesting that cleavage in the loop does not affect the ability of the stem-loop to bind to the SLBP. Presumably the two halves of the RNA are held together by non-covalent interactions after the cleavage has occurred, and binding of the SLBP, determined largely by the stem sequences and the flanking sequences, can still occur. Control experiments demonstrated that the SLBP did not induce cleavage of the RNA in the loop. The bound RNA eluted from the complex was intact (data not shown).

The 3' labeling reaction adds an additional cytosine onto the 3' end of the RNA. In addition there is no longer a 3' hydroxyl at the end of the RNA, but a 3' phosphate from the pCp. The addition of the extra nucleotide and phosphate did not affect the ability of the RNA to bind to either the nuclear or polyribosomal SLBP (not shown). This probe was used to map the 5' boundary of the binding site. The first 6 nt could be removed from the 5' end of the RNA without affecting binding as evidenced by the similar intensity of the bands 3–7 in the bound lanes and the ladder (Fig. 7, lanes 7 and 9). Thus at least 3 nt 5' of the stem-loop are required for efficient binding. There were shorter RNAs which bound to the SLBP and these again map to cleavages in the loop region. These results define a minimum binding site which contains at

Figure 7. Minimum binding site for the SLBP. (A) The wild-type stem-loop sequence was labeled at the 5' end with [γ - 32 P]ATP using polynucleotide kinase. A set of partial cleavage products was generated using KOH. The cleavage products were incubated with a nuclear extract, the SLBP-RNA complex resolved by gel electrophoresis, eluted from the gel and the RNA purified and analyzed on a denaturing gel. Different amounts of the KOH digest used for the mobility shift experiment was analyzed in lanes 1–3 to allow comparison of the intensities of the bound RNAs. The numbers next to the gel refer the sequence of the probe shown below the figure and the internal cleavage sites are indicated by arrows. A partial digest with U2 RNase (lane 5) and with RNase A (lane 7) was analyzed to allow precise identification of the nucleotides in the ladder. The intensities of the cleavage ladder (particularly nucleotides 27–29) was not uniform and the mobility of the fragments was also not uniform, but there was a severe compression in the stem-region between nucleotides 10 and 15. The pattern of the cleavage ladder was very reproducible (cf. lanes 3, 6 and 8) allowing us to make unambiguous assignments based on the cleavage at position A₂₀ at the top of the stem with RNase U2 (lane 5) and the cleavage of the pyrimidines in the loop by RNase A (lane 7). (B) The 30 nt stem-loop probe was labeled at the 3' end with [32 P]pCp using RNA ligase. A set of partial cleavage products was generated using KOH and the mixture was analyzed using the mobility shift assay. The RNA in the specific nuclear complex was eluted from the gel and analyzed by gel electrophoresis under denaturing conditions (lane 1) and analyzed compared to different amounts of the partial cleavage products (lanes 2–4).

Figure 8. Stability of nuclear SLBP-RNA complex. 12 μ g of nuclear proteins from mouse myeloma cells were incubated with the wild-type RNA probe for 20 min. After equilibrium was reached, a 1000-fold molar excess of unlabeled synthetic wt RNA as competitor was added to the reaction and the complexes analyzed after varying amounts of time (1 min to 6 h, given above each lane). The band labeled N represents the nuclear complex, and the band labeled RNA represent free probe. Lane 11 is a reaction in which the 1000-fold molar excess of competitor was mixed with the probe prior to addition of the nuclear extract.

least 3 nt 5' of the stem loop and 2–3 nt 3' of the stem-loop. The loop does not have to be covalently closed for binding. This result is consistent with the results above that expanding the size of the loop did not dramatically reduce binding. Thus it is unlikely that the loop must be in a strictly defined conformation for binding the SLBP. Instead it is likely that the conformation of the stem and flanking nucleotides is the most critical determinant of binding specificity.

The SLBP-RNA complex is very stable *in vitro*

The half-life of the SLBP stem-loop complex *in vitro* was measured by addition of a 1000-fold molar excess of unlabeled competitor to the preformed complex. The complex is stable, with significant complex persisting for at least 6 h after addition of unlabeled competitor (Fig. 8, lane 10). By comparing the relative amount of labeled complex and free RNA in each sample, a half-life of greater than 4 h for the complex was estimated. Similar results were obtained with the nuclear (Fig. 8) and polyribosomal extracts (not shown). Thus once bound, the SLBP remains tightly associated with the RNA. This result is consistent with the possibility that the SLBP binds to the pre-mRNA in the nucleus and remains with the mRNA throughout its lifetime.

DISCUSSION

The 3' end of histone mRNA is involved in a number of different aspects of histone mRNA metabolism, including the processing of the pre-mRNA, transport of the RNA to the cytoplasm (12–14), translation of the mRNA (12) and regulation of the half-life of the RNA (13,15). Since the 3' end is common to all five classes of replication-dependent histone mRNAs, it is an important target for coordinate regulation of the 50–100 different histone mRNAs present in mammalian cells. The 26 nt at the 3' end of histone mRNA has been highly conserved among metazoan histone mRNAs, consistent with an important function for this entire sequence.

It is likely that the function of the histone mRNA 3' end is mediated through interaction with specific factors. A factor present in the nucleus which binds the stem-loop and facilitates

the processing of the histone pre-mRNA has been described but has not been well characterized (7,9). We have previously reported that there are factors which bind to the isolated stem-loop sequence present in both the nucleus and the polyribosomes. As judged by UV-crosslinking the nuclear and polyribosomal complexes share the same 45 kDa polypeptide, although the complexes found in the nucleus and polyribosomes are different in mobility (22). We have suggested that this 45 kDa polypeptide, termed the stem-loop binding protein (SLBP), may be involved in 3' end formation (23) and then accompanies the histone mRNA to the cytoplasm where it is a component of the histone mRNP (27) and plays a role in regulating the half-life of histone mRNA (13).

Previously we have shown that base changes in the loop and changes in the top base pair of the stem alter the binding of the SLBP and affect processing of the histone mRNA *in vivo* (23). In this paper we show that the sequences of the stem and the sequences flanking the stem-loop are also critical for binding of the SLBP and demonstrate that a precise conformation of the loop is not critical for binding.

Sequence requirements for SLBP binding

Many sequence-specific RNA binding proteins bind to single-stranded loops in RNAs and the sequences of the stem are much less important for binding (28). Thus many of the ribosomal proteins (29) and some of the snRNP specific proteins (30–32) bind to conserved single-stranded loops which are formed by stem sequences whose structure, but not sequence, have been conserved. Other RNA binding proteins, the HIV tat protein (33–36), or the iron response element binding protein (37), require bulges in the stem for efficient binding. These proteins recognize specific RNA structures formed as a result of imperfections in the double-stranded RNA.

The SLBP does not recognize a simple primary or secondary structure in the 3' end of histone mRNA but more likely recognizes a structure formed by folding of the highly conserved 26 nt sequence which includes the five nucleotides 5' and five nucleotides 3' of the stem-loop. Strikingly, despite the fact that the size and sequence of the loop has been conserved, alterations which expand the size of the loop by one or two nucleotides did not reduce binding of the SLBP as much as changes in the structure of the stem (Fig. 6). In addition, single cleavages in the loop in the native RNA do not affect the ability of the SLBP to bind (Fig. 7), further supporting the idea that the conformation of the loop is not critical for function of the 3' end. Point mutations in the loop reduce binding of the SLBP, particularly mutations which substitute purines for the conserved uridines at positions 1 and 3 of the loop. It is possible that these mutations disrupt the structure of the RNA rather than interfering directly with specific RNA protein contacts.

The sequence of the flanking regions are also important for binding. The footprinting experiment (Fig. 7) demonstrates that at least three nucleotides 5' of the stem-loop and three nucleotides 3' of the stem-loop are required for efficient binding. Note that in this experiment a reduction in affinity by a factor of five would have been readily detected as a reduced intensity of the bands in the bound fraction, and hence the boundaries of the binding site defined by this experiment are quite stringent. The requirement for flanking sequences is sequence specific since a similar reduction in affinity is caused by deleting the sequences (Fig. 5)

or by replacing them with other sequences (Fig. 6). There was a greater effect of altering or deleting sequences 5' of the stem-loop than of altering the sequences 3' of the stem-loop (Fig. 6). This is best demonstrated by the mutant SL(Δ 5'), which has all the wild-type sequences starting 2 nt prior to the 5' flanking region and bound the SLBP much less well than the SL(Δ 3') mutant which lacks the entire 3' flanking region (Fig. 5).

The SL(GC)₃ mutant has the wild-type flanking regions and the normal stem-loop sequence but the stem has been lengthened to 7 nt. This mutation, with an intact stem-loop, is much more deleterious to SLBP binding than simply adding extra uridines to the loop. Taken together these results suggest that the SLBP binding site is a complex RNA structure formed by the interaction of the flanking sequences with the stem-loop.

Possible function of the SLBP

All of the mutations described previously (22,23) and in this paper showed similar effects on binding of both the nuclear and polyribosomal complexes, consistent with the possibility that both of these complexes contain similar or identical RNA binding components. The SLBP likely mediates most, if not all, the functions of the 3' end of histone mRNA. The initial step in histone mRNA metabolism is formation of the 3' end of the mRNA by endonucleolytic cleavage. The stem-loop is not essential for processing *in vitro* (8,11), although in most histone genes and in extracts from most cells contributes a large extent to the efficiency of processing. However, in transfection experiments, we have shown that the stem-loop sequence is critical for processing *in vivo* (23). Mutations which greatly reduce SLBP binding do not produce any mRNA because the pre-mRNA is degraded before it can be processed (23). The effect of the mutations on processing, measured by accumulation of histone mRNA, parallels the effect on binding of the SLBP (23), suggesting that the SLBP plays a critical role in processing. However, we stress that there is not definitive evidence that nuclear complex we detect by the mobility shift assay is the same as the hairpin-binding factor required for 3' processing. It is possible that the SLBP binds to the histone pre-mRNA shortly after transcription past the stem-loop and participates in 3' end formation, or alternatively the SLBP may not bind until after the processing reaction has been completed.

A striking feature of the SLBP-stem-loop RNA complex is the extreme stability of the binding *in vitro*. Once formed, the complex was not dissociated during exposure to a large excess of competitor stem-loop RNA over a 6 h incubation period (Fig. 8). This is consistent with the possibility that once the SLBP associates with the histone pre-mRNA in the nucleus, it stays with the RNA and accompanies the RNA to the cytoplasm where it is part of the histone mRNP on the polyribosomes. Subsequent degradation of the histone mRNA occurs 3' to 5' on the polyribosomes (38,39). There are two intermediates detected in histone mRNA degradation, one at -5 corresponding to removal of the 3' flanking sequence, and one at -12 corresponding to a cleavage in the loop (40). Each of these intermediates could still have the SLBP bound to the mRNA, even the -12 cleavage which is at a position in the loop which did not affect binding of the SLBP. Thus initial cleavage of the 3' end of histone mRNA could occur with the SLBP still bound to the RNA, as suggested for removal of the polyA tail from polyadenylated mRNAs (41).

Complete degradation of the histone mRNA 3' to 5' (39) may require the subsequent removal of the SLBP.

It is likely that the SLBP plays a critical role in all aspects of the metabolism of histone mRNA and is a common mediator of the coordinate post-transcriptional regulation of histone mRNA levels.

ACKNOWLEDGEMENT

This work was supported by NIH grant GM29832 to WFM.

REFERENCES

- 1 Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) *Cell*, **41**, 349–359.
- 2 Marzluff, W.F. (1992) *Gene Exp.* **2**, 93–97.
- 3 Gick, O., Krämer, A., Keller, W. and Birnstiel, M.L. (1986) *EMBO J.* **5**, 1319–1326.
- 4 Mowry, K.L. and Steitz, J.A. (1987) *Science*, **238**, 1682–1687.
- 5 Bond, U.M., Yario, T.A. and Steitz, J.A. (1991) *Genes Dev.* **5**, 1709–1722.
- 6 Cotten, M., Gick, O., Vasserot, A., Schaffner, G. and Birnstiel, M.L. (1988) *EMBO J.* **7**, 801–808.
- 7 Vasserot, A.P., Schaufele, F.J. and Birnstiel, M.L. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 4345–4349.
- 8 Mowry, K.L., Oh, R. and Steitz, J.A. (1989) *Mol. Cell. Biol.* **9**, 3105–3108.
- 9 Melin, L., Soldati, D., Mital, R., Streit, A. and Schümperli, D. (1992) *EMBO J.* **11**, 691–697.
- 10 Gick, O., Krämer, A., Vasserot, A. and Birnstiel, M.L. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8937–8940.
- 11 Streit, A., Koning, T.W., Soldati, D., Melin, L. and Schümperli, D. (1993) *Nucleic Acids Res.* **21**, 1569–1575.
- 12 Eckner, R., Ellmeier, W. and Birnstiel, M.L. (1991) *EMBO J.* **10**, 3513–3522.
- 13 Williams, A.S., Ingeldue, T.C., Kay, B.K. and Marzluff, W.F. (1994) *Nucleic Acids Res.* **22**, 4660–4666.
- 14 Sun, J.-H., Pilch, D.R. and Marzluff, W.F. (1992) *Nucleic Acids Res.* **20**, 6057–6066.
- 15 Pandey, N.B. and Marzluff, W.F. (1987) *Mol. Cell. Biol.* **7**, 4557–4559.
- 16 Schümperli, D. (1986) *Cell*, **45**, 571–572.
- 17 Marzluff, W.F. and Pandey, N.B. (1988) *Trends Biochem. Sci.* **13**, 49–52.
- 18 Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D. and Marzluff, W.F. (1991) *Mol. Cell. Biol.* **11**, 2416–2424.
- 19 Stauber, C., Lüscher, B., Eckner, R., Lotscher, E. and Schümperli, D. (1986) *EMBO J.* **5**, 3297–3303.
- 20 Stauber, C. and Schümperli, D. (1988) *Nucleic Acids Res.* **16**, 9399–9413.
- 21 Lüscher, B., Stauber, C., Schindler, R. and Schümperli, D. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 4389–4393.
- 22 Pandey, N.B., Sun, J.-H. and Marzluff, W.F. (1991) *Nucleic Acids Res.* **19**, 5653–5659.
- 23 Pandey, N.B., Williams, A.S., Sun, J.-H., Brown, V.D., Bond, U. and Marzluff, W.F. (1994) *Mol. Cell. Biol.* **14**, 1709–1720.
- 24 Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) *Nucleic Acids Res.* **15**, 8783–8794.
- 25 Roberts, S.B., Emmons, S.W. and Childs, G. (1989) *J. Mol. Biol.* **206**, 567–577.
- 26 Scharl, E.C. and Steitz, J.A. (1994) *EMBO J.* **13**, 2432–2440.
- 27 Marzluff, W.F. and Hanson, R.J. (1993) In Belasco, J.G. and Brawerman, G. (eds), *Control of Messenger RNA Stability*. Academic Press, San Diego, pp. 267–290.
- 28 Frankel, A.D., Mattaj, I.W. and Rio, D.C. (1991) *Cell*, **67**, 1041–1046.
- 29 Stern, S., Powers, T., Changchien, L.-M. and Noller, H.F. (1989) *Science*, **244**, 783–790.
- 30 Tsai, D.E., Harper, D.S. and Keene, J.D. (1991) *Nucleic Acids Res.* **19**, 4931–4936.
- 31 Jessen, T.-H., Oubridge, C., Teo, C.H., Pritchard, C. and Nagai, K. (1991) *EMBO J.* **10**, 3447–3456.
- 32 Hall, K.B. (1994) *Biochemistry*, **33**, 10076–10088.
- 33 Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) *Science*, **257**, 76–80.
- 34 Churcher, M.J., Lamont, C., Hamy, F., Dingwall, C., Green, S.M., Lowe, A.D., Butler, P.J.G., Gait, M.J. and Karn, J. (1993) *J. Mol. Biol.* **230**, 90–110.
- 35 Hamy, F., Asseline, U., Grasby, J., Iwai, S., Pritchard, C., Slim, G., Butler, P.J.G., Karn, J. and Gait, M.J. (1993) *J. Mol. Biol.* **230**, 111–123.
- 36 Gait, M.J. and Karn, J. (1993) *Trends Biochem. Sci.* **18**, 255–259.
- 37 Jaffrey, S.R., Haile, D.J., Klausner, R.D. and Harford, J.B. (1993) *Nucleic Acids Res.* **21**, 4627–4631.
- 38 Ross, J., Peltz, S.W., Kobs, G. and Brewer, G. (1986) *Mol. Cell. Biol.* **6**, 4362–4371.
- 39 Ross, J. and Kobs, G. (1986) *J. Mol. Biol.* **188**, 579–593.
- 40 McLaren, R.S. and Ross, J. (1993) *J. Biol. Chem.* **268**, 14637–14644.
- 41 Sachs, A.B. and Deardorff, J.A. (1992) *Cell*, **70**, 961–973.