Electrostatic Contribution of Serine Phosphorylation to the *Drosophila* SLBP−Histone mRNA Complex†

Roopa Thapar,‡ William F. Marzluff,*,‡,§ and Matthew R. Redinbo‡,§,⊥

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

Received December 24, 2003; Revised Manuscript Received April 22, 2004

ABSTRACT: Unlike all other metazoan mRNAs, mRNAs encoding the replication-dependent histones are not polyadenylated but end in a unique 26 nucleotide stem−loop structure. The protein that binds the 3′ end of histone mRNA, the stem−loop binding protein (SLBP), is essential for histone pre-mRNA processing, mRNA translation, and mRNA degradation. Using biochemical, biophysical, and nuclear magnetic resonance (NMR) experiments, we report the first structural insight into the mechanism of SLBP−RNA recognition. In the absence of RNA, phosphorylated and unphosphorylated forms of the RNA binding and processing domain (RPD) of *Drosophila* SLBP (dSLBP) possess helical secondary structure but no well-defined tertiary fold. *Drosophila* SLBP is phosphorylated at four out of five potential serine or threonine sites in the sequence DTAKDSNSDSDSD at the extreme C-terminus, and phosphorylation at these sites is necessary for histone pre-mRNA processing. Here, we provide NMR evidence for serine phosphorylation of the C-terminus using ^{31}P direct-detect experiments and show that both serine phosphorylation and RNA binding are necessary for proper folding of the RPD. The electrostatic effect of protein phosphorylation can be partially mimicked by a mutant form of SLBP wherein four C-terminal serines are replaced with glutamic acids. Hence, both RNA binding and protein phosphorylation are necessary for stabilization of the SLBP RPD.

Most eukaryotic mRNAs are posttranscriptionally modified by the addition of a 5′ 7-methyl-guanine cap and a 3′ poly(A) tail. These modifications are necessary for the proper control of mRNA metabolism, which requires the assembly of large complexes that regulate translation initiation, mRNA stability, and mRNA degradation. Replication-dependent histone genes encode the only known family of mRNAs that are not polyadenylated. Unlike other eukaryotic mRNAs that end in a polyA sequence, replication-dependent histone mRNAs end in a highly conserved, 26 nucleotide stem−loop structure. The stem−loop of replication-dependent histone mRNAs is involved in pre-mRNA processing, translation, and stability (*1*, *2*). Mature histone mRNAs are generated by endonucleolytic cleavage four or five nucleotides downstream of the stem−loop to form the 3′ end. The first step in this processing reaction involves the binding of a 32 kDa stem−loop binding protein (SLBP)[1] to the stem−loop and recruitment of the U7 small nuclear ribonucleoprotein (snRNP) to a site downstream of the stem−loop in the pre-mRNA (*3*). The pre-mRNA is subsequently cleaved, and the mature mRNA is exported to the cytoplasm. SLBP is also a component of the cytoplasmic histone messenger ribonucleoprotein (mRNP) complex (*4*). The high-affinity ($K_D \approx 1$ nM) interaction (*5*, *6*) between SLBP and the mRNA stem−loop is crucial for efficient processing, translation, and regulation of degradation of histone mRNA in mammalian cells.

Recent biochemical studies (*3*−*6*) on the SLBP−RNA complex and solution NMR studies (*7*, *8*) on the free RNA demonstrate that the mode of interaction of SLBP with histone stem−loop RNA is likely to be unique. The solution NMR structures of 24-nt and 28-nt stem−loops confirm that the RNA forms a canonical six base-pair A-form stem that starts with the highly conserved GC base pair at the base and ends in the UA base pair. The UUUC tetraloop is stabilized by base stacking interactions. Few NOEs are observed for bases flanking the stem, and NMR relaxation measurements indicate that these bases are disordered in solution. Mutagenesis experiments (*3*, *5*, *6*) on the RNA reveal that the SLBP−RNA interaction involves sequence-specific contacts among the mRNA stem, the loop, and the flexible flanking sequences, particularly the 5′ flanking

† Supported by National Institutes of Health (NIH) Research Grant GM58961 to W.F.M, an NIH Supplement Grant GM29832 to W.F.M. and R.T., NIH Research Grant CA90604 to M.R.R. and a Burroughs Wellcome Career Award to M.R.R.

* To whom correspondence should be addressed. E-mail address: marzluff@med.unc.edu. Phone: (919) 962-2141. Fax: (919) 962-1274.

‡ Department of Biochemistry and Biophysics.

§ Program in Molecular Biology and Biotechnology.

⊥ Department of Chemistry.

[1] Abbreviations: SLBP, stem−loop binding protein; dSLBP, *Drosophila* stem−loop binding protein; RPD, RNA binding and processing domain; RBD, RNA binding domain; HSQC, heteronuclear single quantum coherence; PABP, poly(A)-binding protein; 3′-UTR, 3′ untranslated region; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; MALDI, matrix-assisted laser desorption ionization; PFG, pulsed-field gradient; PG-SLED, pulsed gradient stimulated echo longitudinal encode−decode; EMSA, electrophoretic mobility shift assay; CD, circular dichroism; CREB, cyclic AMP-response element-binding protein; CBP, CREB-binding protein; KIX, CREB-binding domain of CBP; KID, kinase-inducible activation domain; c-Myb, c-myb protooncogene product.

10.1021/bi036315j CCC: $27.50
Published on Web 07/02/2004

sequence. Mutation of both the first and third uridine in the loop and base transversions in the stem at the second GC base pair decrease binding affinity >200-fold, whereas mutation of the sixth UA base pair to a UG base pair reduces affinity by ~15-fold (*9*). In contrast to known RNA binding domains, such as the KH domain and the RRM domain that recognize non-Watson−Crick or exposed bases in single-stranded loops and bulges or the dsRBD that recognizes the double-stranded stem in the minor groove without much sequence specificity, SLBP appears to make base-specific contacts with the single-stranded flanking and loop regions, as well as the double-stranded stem in a sequence-specific manner. It is not clear how this small 70 amino acid RNA binding domain (RBD) that bears no homology to any other known protein achieves this extensive mode of RNA interaction.

Drosophila SLBP (dSLBP) is phosphorylated at multiple sites in the N- and C-terminal domains (*10*, *11*). Several serine and threonine phosphorylation sites have been identified in the RNA binding and processing domain. Electrospray ionization mass spectrometric analysis of C-terminal deletion mutants also suggests that there are four (out of five possible) phosphorylation sites in the extreme C-terminus of dSLBP in the sequence DTAKDSNSDSDSD (*10*). Phosphorylation of the extreme C-terminus is necessary for efficient processing of the histone pre-mRNA to the mature mRNA but not for RNA binding in electrophoretic mobility shift assays (EMSA).

We report herein our first insight into the structural changes that occur when both the phosphorylated (P-RBD) and unphosphorylated (RBD) SLBP RBDs bind histone mRNA. Our data suggest that both the SLBP RBD and P-RBD lack a well-defined tertiary fold and fluctuate between an ensemble of partly folded helical conformations in solution. Although spectral changes are observed in response to RNA binding, the presence of RNA is not sufficient to stabilize a single conformation in the unphosphorylated protein. We provide NMR evidence that the C-terminal serines are phosphorylated and show that the role of serine phosphorylation is to increase further the stability of the SLBP−RNA complex by a favorable electrostatic interaction most likely with one or more basic residues in this domain. The effect of phosphorylation can be mimicked partially by mutation of the four C-terminal serines to glutamic acids in a mutant that we term 4E-RPD. Intriguingly, sulfate anions can also induce folding of the RPD and the presence of both sulfate anions and RNA results in complete folding of the RPD. Our results are consistent with previous biological studies reported for the SLBP−RNA complex and form the basis for ongoing structural studies.

EXPERIMENTAL PROCEDURES

Limited Proteolysis of Baculovirus dSLBP. Full-length, baculovirus-expressed human and *Drosophila* SLBPs were incubated with the proteases EndoLysC, trypsin, Asp-N, and Glu-C at 37 °C at a molar ratio of SLBP/protease of 1:100 at varying time points up to 2 h. Human SLBP was completely degraded within 15 min, whereas the Drosophila isoform gave distinct bands in all reactions. All reactions were stopped at varying time points by the addition of Pefaloc (Boehringer Mannheim) and freezing the reaction on dry ice. SDS−PAGE loading buffer was added to the reactions, which were boiled for 10 min, and then the samples were analyzed by 14% SDS−PAGE. The Endo-Lys C digest gave only two bands, which were subjected to detailed analysis by MALDI mass spectrometry. The fragments were gel-eluted, and a precise whole mass was obtained on each fragment. In addition, the two fragments were subject to total tryptic digest, the resulting peptides were isolated by HPLC, and masses were obtained on the individual peptides. The total coverage of peptides mapped to the C-terminal RPD domain was 60%, whereas that for the N-terminal domain was 40%. Peptides for which masses were not obtained lay between sequence segments that were unambiguously identified using this approach, thereby increasing our confidence level in the right domain cutoffs. Truncation of the fragments identified by mass spectrometry resulted in severe solubility and aggregation problems, providing further evidence that our fragments represented structural domains in the intact protein. These analyses were performed by Dr. Christoph Borchers' group (Department of Biochemistry and Biophysics, UNC-CH), and the analysis of the phosphorylated state of the baculovirus-expressed RPD has been reported elsewhere (*10*).

Sample Preparation. The C-terminal 105 amino acids of the dSLBP, termed the RPD (residues 172−276), were subcloned into the *Nde*I and *Eco*RI restriction sites of the bacterial and baculoviral expression vectors pET21a (Novagen) and pFAST BacHTa (Invitrogen), respectively. Unlabeled and uniformly ^{15}N-labeled SLBP RPD was expressed from the vector pET21a by growing BL21(DE3) RIL Codon Plus cells (Stratagene) transformed with pET21a−RPD at 20 °C in the presence of 1 mM IPTG in Luria broth or minimal media, respectively. Under these conditions, about 50% of the protein was recovered from the soluble fraction. The RPD was purified using cation-exchange chromatography over an S-sepharose column (Amersham Pharmacia Biotech) followed by gel filtration chromatography using an S100 column (Amersham Pharmacia Biotech). The proteins were >95% pure as judged by SDS−PAGE. Phosphorylated forms of the RPD and full-length SLBP proteins were obtained by infecting Sf9 cells with RPD−baculovirus and growing the Sf9 cells in Grace's medium for 48 h. Baculovirus-expressed, full-length RPD proteins had a 20 residue N-terminal tag that encoded a His tag and were hence purified using nickel affinity chromatography. The dSLBP RPD-4E mutant in which all four C-terminal serines were mutated to glutamic acids using Quickchange mutagenesis was expressed in the vector pET28a (Novagen) with a C-terminal His tag and purified using immobilized metal affinity chromatography (IMAC) over a Ni^{2+} column. Unlike the RPD, which was mostly insoluble when expressed with a His tag, the His-tagged 4E mutant was quite soluble when expressed at 20 °C. The identity of all constructs was confirmed by DNA sequencing as well as electrospray mass spectrometery. The 28-mer RNA (5′GGCCAAAGGCCCUUUUCAGGGCCACC CA3′) used in our studies corresponds to the sequence of the mammalian histone H4 hairpin for which NMR structures have previously been described (*7*, *8*). The RNA was custom synthesized, deprotected, and PAGE-purified by Dharmacon, Inc. To ensure that >95% of the RNA was in the stem−loop form, the lyophilized RNA was taken up in NMR buffer, heated to 95 °C for 10 min, and then snap-cooled on ice.

The protein−RNA complex used for NMR studies was formed by first exchanging the protein into 20 mM Tris, pH 7.0, 50 mM KCl, 0.2 mM EDTA, 10% D_2O, and 0.1% sodium azide and then adding a suitable volume of a highly concentrated RNA solution to the protein sample to form the SLBP RPD−RNA complex.

NMR Measurements. NMR data were collected at 25 °C and pH 7.0 on an Inova 600 MHz spectrometer equipped with either a 5 mm z-gradient triple resonance probe for collection of HSQC or NOESY data or a 5 mm broadband probe used for ^{31}P studies. 2D (^{15}N,^{1}H) HSQC spectra were collected on the free and RNA-bound forms of ^{15}N-labeled RPD. A 3D (^{15}N,^{1}H) NOESY−HSQC (τ_m = 200 ms) was collected with a 1 mM sample of dSLBP RPD with 16 transients per scan and 200 increments in t_1 and 64 increments in t_2. Spectra were processed using the program Felix 98.2 (Accelrys).

Pulsed-Field Gradient (PFG) NMR. Diffusion experiments were performed using a PG-SLED (pulsed gradient stimulated echo longitudinal encode−decode) sequence that has been previously described (*12*). Typically, a series of 15−20 1D proton spectra were acquired by varying the strength of the diffusion gradient between 100% and 10% of the maximum peak intensity using 64 transients per spectrum and a 1 mM unlabeled protein sample. A concentration of 1 mM is routinely used in these experiments to obtain sufficient S/N at longer gradient strengths. NMR spectra were processed and the methyl proton peak intensities were integrated using the software package VNMR (Varian, Inc.). Peak intensities, $s(g)$, were fit as a function of gradient strength (g) using the equation $s(g) = A\ e^{-\delta g^2}$ to obtain the observed decay rate (d). The value of the protein hydrodynamic radius, R_h^{prot}, was calculated using dioxane as a reference molecule, such that the ratio $R_h^{prot}/R_h^{ref} = d_{ref}/d_{prot}$. The effective hydrodynamic radius used for dioxane (R_h^{ref}) was 2.12 Å (*12*). The theoretically predicted values of the hydrodynamic radii were calculated by the method of Uversky (*13*).

Circular Dichroism. All experiments were performed in 10 mM potassium phosphate buffer, pH 7.0, at 20 °C. The conformational stability of the RNA binding domains in the presence and absence of the RNA was determined using thermal denaturation. Thermal unfolding was monitored by observing the change in ellipticity at 222 nm as a function of increasing temperature. The heating rate was 30 °C per hour in a cuvette with a path length of 1 mm.

Sedimentation Equilibrium. Sedimentation equilibrium measurements for RPD and P-RPD were performed on 0.3 mM samples in NMR buffer described above in a Beckman XL-A analytical ultracentrifuge using a Ti60 four-hole rotor. The rotor speed was set at 20 000 rpm for 18 h and then set to 45 000 rpm for an additional 8 h. At the end of this period, the meniscus was depleted as could be ascertained by comparison of the last absorption profiles. Absorbance scans were recorded at 2 h intervals at 295 nm and 25 °C. Data analysis was performed using the software program Origin XL-A.

Calculation of Phosphate pK_a Values. A series of 1D ^{31}P NMR spectra were collected at different pH values on an Inova 500 MHz spectrometer at 25 °C. The pK_a values were determined from a least-squares fit of the ^{31}P chemical shift as a function of pH to the equation $\delta = [\delta_2(10^{pH-pK_a}) + \delta_1]/[1 + 10^{pH-pK_a}]$, where δ_2 and δ_1 represent the chemical shifts of the dianionic and monoanionic forms of the phosphate group, respectively. The random-coil ^{31}P-serine, ^{31}P-threonine, and ^{31}P-tyrosine chemical shifts were obtained from previous studies as mentioned in Table 1 (*14*, *15*).

Measurement of Off Rates. Since the off rate of the protein−RNA complex is very slow, that is $<10^{-2}$ min^{-1} at 37 °C, it is not possible to use BIACORE to accurately determine the dissociation constant. Hence, the dissociation rate constant (or off rate) for the 28-mer RNA for the phospho-RPD (P-RPD) and phosphorylated full-length SLBP were measured using EMSA assays by monitoring the decrease in intensity of the bound complex over 2 days when run on a native gel. The 28-mer RNA used in these studies was 5′-end labeled with ^{32}P-ATP by incubating the RNA with T4 kinase and ^{32}P-γ-ATP for 2 h at 37 °C. Protein (final concentration 2 μM) samples were mixed with ^{32}P-labeled RNA hairpin (final concentration 150 nM) in a total volume of 50 μL and kept on ice for 60 min to preequilibrate the protein−^{32}P-RNA complex. Unlabeled RNA (final concentration 15 μM) was added to the complex, and the complex was immediately placed in a 37 °C water bath to facilitate dissociation. Five microliter samples were removed from the reaction over 2 days, snap-cooled on dry ice, and then immediately placed at −80 °C. The samples were analyzed on a native gel and the bound complex was quantified using a PhosphorImager. The off rates were obtained by fitting the intensity of the fraction [RNA] bound vs time to a single-exponential decay of the form $y = C_t[\exp(-k_{off}t)] + C$, where C_t is the intensity of the bound complex at time t.

RESULTS

Identification of the Minimal RNA Binding Domain (RBD). The sequence of the conserved 26 nucleotides at the 3′ end of histone mRNA is shown in Figure 1A. The RBD of SLBP bears no sequence homology to any known RNA binding protein, and the protein is predicted to have low sequence complexity in threading databases. The RBD has been conserved in evolution, and the sequence alignment of the RBD from mammalian SLBP (*16*), the two SLBPs from frogs (*4*), *Drosophila* SLBP (*17*) and *Caenorhabditis elegans* SLBP (*18*) is shown in Figure 1B. Efforts to predict the overall fold of human SLBP RBD (hRBD) using computational genomics predicted a domain that is largely all α-helical (Samudrala, R., University of Washington, personal communication). Mutational analysis of conserved basic and aromatic residues the hRBD has identified some of the residues that may be important for RNA binding and for pre-mRNA processing (*3*) indicating that both electrostatic and base stacking interactions are important for RNA recognition. In particular, two arginines (Arg197 and Arg198, dSLBP numbering, Figure 1), two lysines (Lys200 and Lys206), and two aromatic residues (Tyr211, Tyr214) are required for efficient binding of SLBP to the stem−loop RNA structure (*3*).

To identify a minimal RNA binding domain that is stably folded and amenable to structural studies, we subjected human (hSLBP) and dSLBP to limited proteolysis (Figure 2). Under limiting conditions, the protease EndoLysC cleaved dSLBP yielding two distinct fragments. In contrast, human SLBP was almost completely degraded, suggesting that

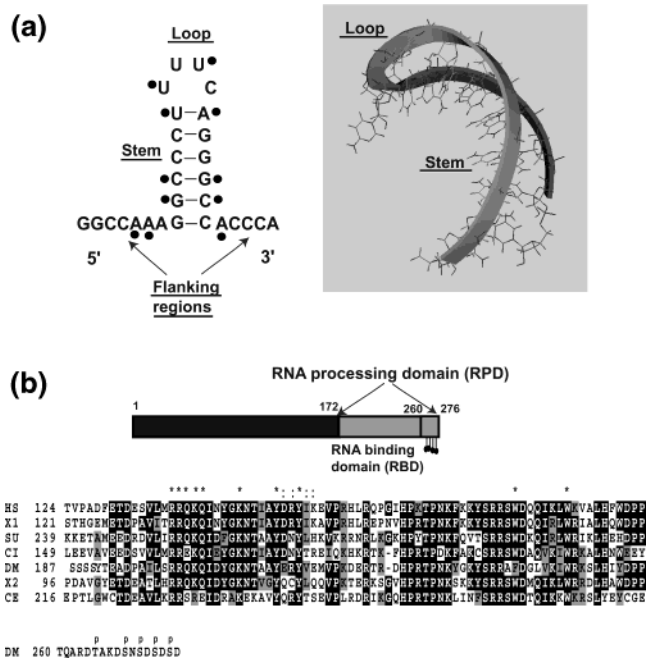

FIGURE 1: The sequence of the stem−loop at the 3′ end of histone mRNA and the RBD of SLBP. Panel A shows the base sequence and three-dimensional structure of the 28 nucleotide hairpin structure containing a six base-pair stem and a four nucleotide tetraloop. Nucleotides that are important for sequence specific recognition of SLBP are highlighted in black circles. Panel B shows the sequence alignment of SLBP from human (HS), *Xenopus* (X1 and X2), *D. melanogaster* (DM), *C. elegans* (CE), sea urchin (SU), and *Ciona* (CI). Amino acids that are required for RNA binding based upon mutagenesis studies are indicated by (∗), whereas amino acids that are required for 3′-end processing in hSLBP are shown as (:). The potential sites of phosphorylation at the extreme C-terminus of the *Drosophila* protein that are important for histone pre-mRNA processing are indicated with a "p". The alignment was generated in ClustalX, and the figure was made using BOXSHADE. Residues that are identical are highlighted in black, whereas residues that are homologous but not identical are highlighted in gray.

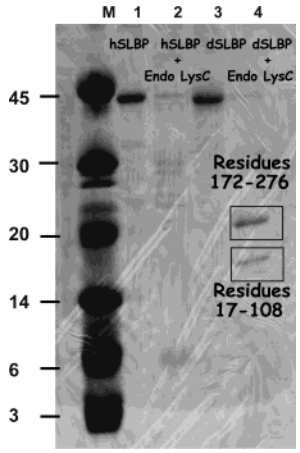

FIGURE 2: Limited proteolysis of full-length SLBP from human and *Drosophila.* Baculovirus-expressed, full-length proteins were subjected to limited proteolysis with the protease EndoLysC. Full-length human SLBP (lane 1) is rapidly degraded upon cleavage with EndoLysC (lane 2), whereas the dSLBP (lane 3) is cleaved into two fragments (lane 4). The two fragments from dSLBP corresponded to residues 17−108 and 172−276 from the N- and C-terminal domains, respectively. Both full-length SLBP and the RPD migrate more slowly than expected on SDS−PAGE. Lane M is molecular weight markers.

fragments of hSLBP are less stable and not likely to have a well-defined overall fold. The two EndoLysC fragments of dSLBP were subjected to analysis by MALDI mass spectrometry. Masses were obtained both for the intact fragments and for a complete tryptic digest of the fragments. This analysis showed that the two fragments in the EndoLysC digest corresponded to the N- (residues 17−108) and C-terminal fragments (residues 172−276) of dSLBP, respectively. The C-terminal fragment contained the complete RNA binding domain, as well as the additional region from the extreme C-terminus important for RNA processing (*10*). Intriguingly, there are seven additional lysines in this fragment, but the protease EndoLysC did not cleave at these sites indicating that these lysines must lie in a structured environment that is resistant to protease.

Residues 172−276 were expressed in bacterial and baculoviral expression vectors for further biochemical and biological analysis of this domain (*10*). Unlike the mammalian and frog SLBPs, the RPD fragment of dSLBP was readily expressed at high levels as a soluble protein in *Escherichia coli*. Efforts to express smaller fragments such as the minimal RNA binding domain or a protein that was truncated from the C-terminus yielded poor expression levels or insoluble protein in bacteria also suggesting that the domain that we have identified using limited proteolysis and mass spectrometry is the minimal structural domain. Since the 105 amino acid domain identified by limited proteolysis is fully active in RNA binding and histone pre-mRNA processing in vitro (*10*), we refer to it as the RNA processing domain (RPD).

Previous electrospray mass spectrometric analysis of the baculovirus-expressed RPD revealed that the protein was quantitatively phosphorylated at four out of five possible sites, and mass spectrometric analysis of deletion mutants indicated that these are most likely the serines in the sequence DTAKDSNSDSDSD at the extreme C-terminus of the molecule (*10*). Efforts to identify the precise sites of protein phosphorylation by mass spectrometry have so far been unsuccessful since the negatively charged phosphopeptide does not ionize well in the mass spectrometer. As would be expected, the bacterial protein is not phosphorylated. The phosphorylated RPD (P-RPD) is active in histone pre-mRNA processing, whereas the dephosphorylated form is not (*10*). On the basis of the mobility in SDS−PAGE, *Drosophila* SLBP is also phosphorylated in vivo, and this is likely to be the functionally relevant form of the protein (*10*).

To test whether the bacterially expressed protein was functional in RNA binding, we compared the association of dSLBP RPD with 20 nucleotide, 24 nucleotide, and 28 nucleotide RNA hairpins using an electrophoretic mobility shift assay (EMSA). The 28-mer bound the protein most strongly, binding of the 24-mer was slightly reduced, and the 20-mer bound SLBP weakly (Figure 3A). We compared the dissociation rate of baculovirus-expressed phosphorylated full-length SLBP and the P-RPD from the 28 nucleotide hairpin. The off rate for dissociation of a preformed complex was measured in EMSA assays by competing the ^{32}P-labeled RNA with a 100-fold molar excess of unlabeled 28-mer hairpin. The k_{off} for full-length phosphorylated SLBP was estimated to be 7.1×10^{-4} min^{-1}, whereas that of the isolated P-RPD was 3.8×10^{-4} min^{-1} at 37 °C, indicating that the isolated P-RPD has a similar dissociation rate as the full-length protein (Figure 3B). Our data confirms that the isolated P-RPD is a good functional mimic of full-length SLBP and

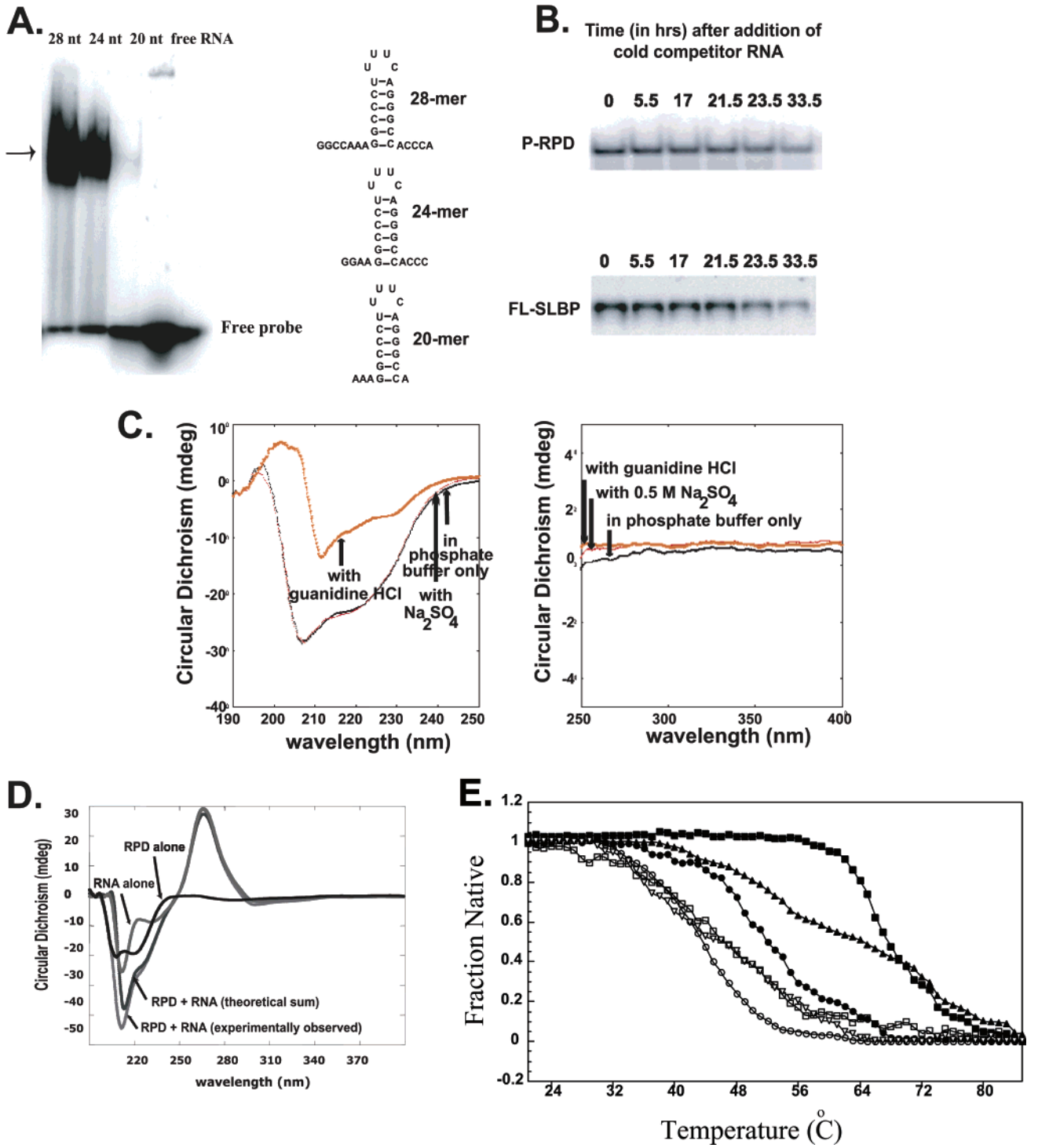

FIGURE 3: Properties of the C-terminal RPD. In panel A, the baculovirus dSLBP RPD was incubated with a radiolabeled stem−loop probe, and after 20 min, the complex was resolved by gel electrophoresis. The position of the complex is indicated with an arrow. Panel B shows off rates for dissociation of baculovirus-expressed phospho-RPD (top) and full-length SLBP (bottom) measured using electrophoretic gel mobility assays. The RNA complexes were kept on ice for 20 min, and then a 100-fold excess of unlabeled RNA was added for the indicated times before loading the reactions on a 8% polyacrylamide gel. A significant amount of protein remains bound to the ^{32}P-labeled RNA after 48 h reflecting the tight affinity of the complex. In panel C, CD spectra of the bacterially expressed RPD were collected in 10 mM potassium phosphate buffer, pH 7.0, at 25 °C with or without the addition of 0.5 M sodium sulfate or 0.4 M guanidine HCl as indicated. The near-UV CD spectrum is shown on the right and the far-UV CD spectrum is on the left. Panel D shows the circular dichroism spectra of unphosphorylated SLBP RPD alone and RNA alone, the experimental CD spectrum of unphosphorylated SLBP RPD complexed to RNA, and the theoretical summed spectrum of the RPD−RNA complex. Comparison of the summed and measured spectra of the complex indicate that no major conformational changes occur in the RNA upon complex formation. RNA binding may be accompanied by small rearrangements or changes in helical structure in SLBP. Panel E shows thermal unfolding transitions for phosphorylated baculovirus P-RPD SLBP (□,■), the bacterially expressed nonphosphorylated RPD SLBP (○,●), and the 4E-RPD SLBP mutant (△,▲) free (open symbols) and complexed to RNA (filled symbols). The RNA-free and -bound forms of SLBP RPD were monitored using circular dichroism at 222 nm. The protein concentrations were 25 μM, and an equimolar ratio of RNA was added to the complexes. The buffer used was 20 mM potassium phosphate, pH 7.0, 0.1 mM EDTA. The data were not fit to a two state unfolding transition since the proteins are not fully folded at pH 7.0 and 25 °C.

that RNA binding of the isolated RPD is likely to mimic full-length SLBP.

The SLBP RPD Is Not Stably Folded in the Absence of RNA. Biophysical characterization of dSLBP RPD indicated that in the absence of RNA the RPD is not stably folded and exchanges between different conformations on the NMR time scale. The far-UV circular dichroism (CD) spectrum (Figure 3C) of dSLBP RPD has double minima at 209 and 222 nm, consistent with the presence of helical structure in the protein. The spectrum lacks positive ellipticity below 200 nm mostly likely reflecting contributions from random coil. The protein also has no near-UV CD signal, which suggests a lack of packing interactions in the core (Figure 3C). The addition of 0.5 M sodium sulfate, which is known to stabilize folded states in disordered proteins such as the protein component of ribonuclease P (*19*) had no effect on the CD spectrum of SLBP RPD (Figure 3C), although it does affect the NMR spectrum of the 4E-RPD (as discussed below), whereas the addition of a small amount of guanidine HCl (0.4 M) resulted in a dramatic loss of helical signal suggesting that the domain does not exist in a well-defined, stable conformation. Consistent with this, both the free RPD and the free P-RPD exhibit a broad thermal unfolding transition (Figure 3E) that may reflect either the loss of a molten globule-like state or loss of helical secondary structure.

Our observations using CD are further substantiated by the (^{1}H,^{15}N) HSQC spectrum of the SLBP RPD (Figure 4A). The dSLBP RPD is soluble up to concentrations of 1 mM making it amenable to study by solution NMR techniques. The (^{1}H,^{15}N) HSQC spectrum shows that the chemical shift dispersion in the amide proton region is no greater than 1 ppm and that most resonances adopt random-coil ^{1}H chemical shifts (8.0−8.4 ppm). Similar to the collapse in backbone NH chemical shifts, we observe degeneracy in the chemical shifts of the side chain NH_2 resonances of glutamine and asparagine residues. A collapse in amide proton chemical shifts is observed in (^{1}H,^{15}N) HSQC spectra under conditions wherein (i) the protein is either unfolded or disordered, (ii) the protein aggregates under NMR conditions, or (iii) the protein exchanges rapidly among a number of conformations (distinct folded states or exchange between folded and unfolded conformations) on the intermediate time scale.

To distinguish among these possibilities and ascertain the nature of the free state, it was imperative to determine whether the protein was aggregated under the NMR conditions used. We performed sedimentation equilibrium experiments on unphosphorylated, bacterially expressed dSLBP RPD and on the phosphorylated, baculovirus-expressed dSLBP P-RPD at concentrations similar to those used for NMR (0.1−0.3 mM). The equilibrium data fit well to a model describing a single species, and the measured molecular weight from sedimentation equilibrium agreed well with the calculated mass for this domain (Figure 5). This result is also supported by elution on a gel filtration column, which shows that the SLBP RPD is predominantly monomeric at NMR concentrations (not shown). Therefore the spectral features of the dSLBP RPD are not due to aggregation. In contrast to the NMR spectra of the N-terminal domain (see accompanying paper), which also shows a collapse in amide proton chemical shifts but gives distinct cross-peaks and is largely unfolded, the cross-peaks in the RPD HSQC are not well-defined suggesting that unlike the N-terminal domain the RPD does not have large segments of unfolded structure. Consistent with this interpretation, a 3D ^{15}N-edited NOESY−HSQC spectrum on SLBP RPD (mixing time 200 ms) is characterized by weak short-range NOEs, such as dNNs, and few long-range NOEs are observed. The NOE data are sparse, consistent with the lack of well-defined tertiary structure in the molecule. However, we do not observe H^N, $H^N(i,i+1)$ and H^α, $H^N(i,i+1)$ NOE cross-peaks, which would be expected if the molecule were completely unfolded or denatured. We conclude therefore that most residues exchange between multiple partly folded conformations in solution.

The hydrodynamic radius of dSLBP RPD was measured using pulsed-field gradient NMR (PFG-NMR, Figure 6), and the measured values were compared to those calculated as described in Figure 6 and Table 1. PFG-NMR can be used to determine the translational diffusion coefficient, D, of a macromolecule. Since D is inversely proportional to the hydrodynamic radius R_H, the hydrodynamic radius can easily be calculated given the values of D and R_H for an internal standard such as dioxane (*12*). Figure 6 shows the decrease in signal intensity of the aliphatic resonances as a function of increasing gradient strength and fit of the data to derive the translational diffusion coefficient. The measured and calculated values of the hydrodynamic radii are summarized in Table 1. The measured values for the SLBP RPD and P-RPD are similar and lie between those calculated for a natively folded and a denatured protein. The PFG-NMR data are consistent with the other NMR and CD measurements suggesting that the 105 amino acid domain does not adopt a stable fold in the absence of RNA, but also does not form large aggregates.

RNA Binding Is Not Sufficient To Stabilize the Unphosphorylated RPD. There are at least two RNA binding proteins, the bacteriophage λ N protein (*20*), which is a transcriptional antiterminator, and the M domain of Ffh (*21*), a component of the signal recognition particle (SRP) in *E. coli*, that have been reported to be completely unfolded in the free state and are stabilized upon binding to RNA. To ascertain whether the SLBP RPD is also stabilized upon RNA binding, we titrated the 28-mer histone mRNA into a solution of RPD and collected CD and NMR data. The CD spectrum (Figure 3D) of a 1:1 dSLBP RPD/RNA complex is very close to the theoretically calculated spectrum obtained by summing the spectra of the RNA and RPD alone, indicating that RNA recognition does not involve major conformational changes. The ellipticity at 222 nm is more diagnostic of helical character in the protein, and a small increase in helical content is observed in the RNA-bound form. Similarly the CD spectrum of the RNA between 250 and 320 nm shows a small decrease in ellipicity, although it retains the overall A-form character. Therefore small but significant conformational rearrangements are expected to occur in both the protein and RNA upon complex formation. When the conformational stability of the RPD was monitored in the presence of RNA using circular dichroism (Figure 3E), the RPD exhibited a T_m that was at least 7 °C higher than that of the free protein. Given the slow off rate of the P-RPD for the RNA and that there is minimal contribution to the CD ellipticity at 222 nm from the RNA itself, the increased stabilization of SLBP must be due to its interaction with RNA.

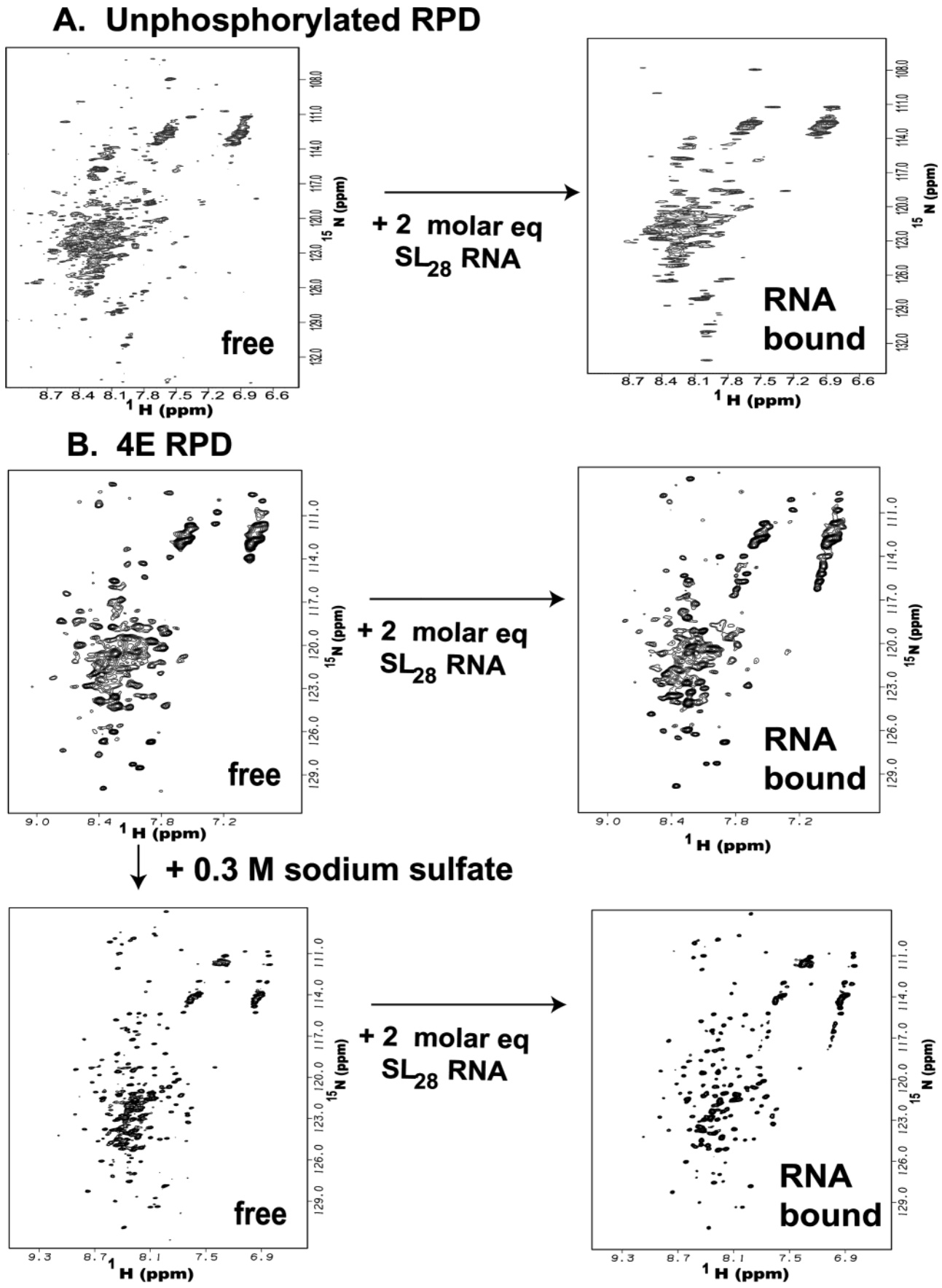

FIGURE 4: $^{15}N,^{1}H-HSQC$ spectra of (A) bacterially expressed unphosphorylated dSLBP RPD (residues 172−276) free and bound to 28-nucleotide stem−loop RNA and (B) dSLBP 4E-RPD free, free but in the presence of 0.3 M sodium sulfate, bound to a 28 nucleotide stem−loop mRNA in a 1:2 (RBD/RNA) complex but in the absence of sulfate, and RNA-bound and sulfated. All spectra were collected at pH 6.9 and 25 °C on a 600 MHz Varian Inova spectrometer. The sample concentration for the unphosphorylated sample in panel A was 0.3 mM, and that for the 4E-RPD sample was 0.5 mM.

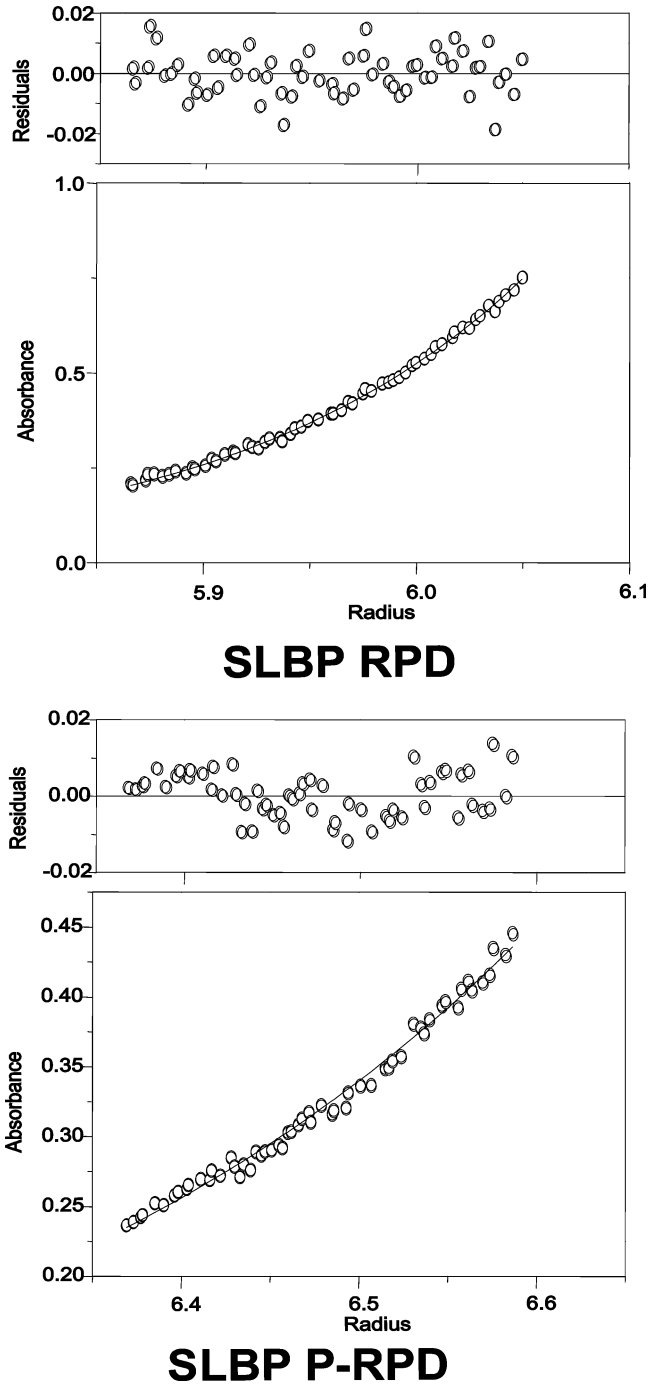

FIGURE 5: Analytical ultracentrifugation data. Absorbance distribution profiles of both phosphorylated and nonphosphorylated forms of the SLBP RPD were measured at 295 nm to allow detection of the protein at a concentration of 0.3 mM. The sedimentation equilibrium data was collected at 20 °C at a rotor speed of 20 000 rpm as described in Materials and Methods and can be modeled as a monomeric species. The theoretical fit to the experimental data is shown as a solid line.

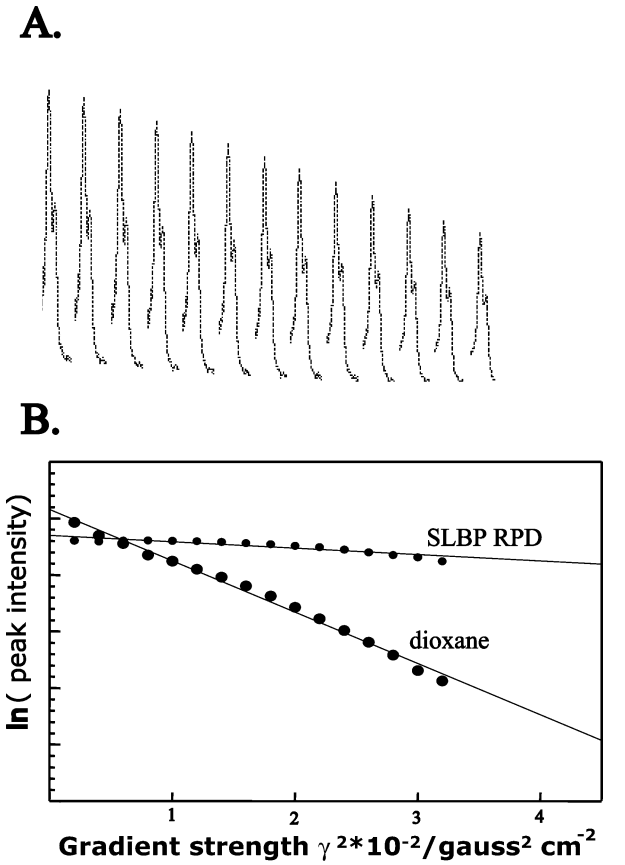

FIGURE 6: Pulsed-field gradient NMR spectroscopy was used to determine the hydrodynamic radius of the bacterially expressed dSLBP RBD and the baculovirus-expressed, phosphorylated dSLBP RBD, In panel A, aliphatic resonances of dSLBP RPD are shown arrayed as a function of increasing gradient strength. In panel B, peak integrals of the aliphatic resonances (1.2−3.0 ppm) were plotted against gradient strength, and the curves were fit as described in Materials and Methods. The slope of the curve is the diffusion coefficient *D*. The hydrodynamic radius was calculated using the radius of dioxane as a standard. The measured and calculated hydrodynamic radii are summarized in Table 1.

Table 1: Experimental and Calculated Hydrodynamic Radii of RNA Binding Domains of dSLBP at 25 °C and pH 7.0[a]

	calcd R_H (Å)		measd R_H (Å)
sample	native state	denatured state	native state
P-RPD	17.8−18.2	31.02−31.8	25.6
RPD	17.9−18.2	31.02−32.2	26.4

[a] The range of calculated R_H values were obtained from that predicted from either pulsed-field gradient NMR according to the relationship proposed by Wilkins et al. (*12*), $R_H = (4.75 \pm 1.1)N^{0.29\pm0.02}$ for a natively folded protein and $R_H = (2.21 \pm 1.07)N^{0.57\pm0.02}$ for a highly denatured polypeptide where *N* is the number of residues in the protein, or that calculated by the method of Uversky (*13*) using gel filtration data, $\log(R_H) = -(0.254 \pm 0.002) + (0.369 \pm 0.001)\log(MW)$ for native proteins and $\log(R_H) = -(0.543 \pm 0.004) + (0.502 \pm 0.001)\log(MW)$ where MW is the average molecular weight of the protein.

The (^{1}H,^{15}N)-HSQC spectrum of ^{15}N-labeled unphosphorylated SLBP RPD complexed with 28-mer RNA at a 1:2 (protein/RNA) molar ratio is shown in Figure 4A. When compared to the HSQC spectrum of the free RPD, the HSQC spectrum of the RNA-bound form of SLBP RPD shows more defined cross-peaks, and the peaks are better resolved. However, there are still many peaks that are either of weak intensity or exchange-broadened in the presence of the RNA. These exchange-broadened peaks do not increase in intensity upon the addition of RNA and most likely reflect the motional properties of the protein in the RNA complex. In addition, the lack of good spectral resolution in the amide proton region in the bound state is consistent with the protein being predominantly helical when bound to RNA. These data indicate that although RNA binding does afford some degree of stabilization to the SLBP RPD, it is not sufficient to

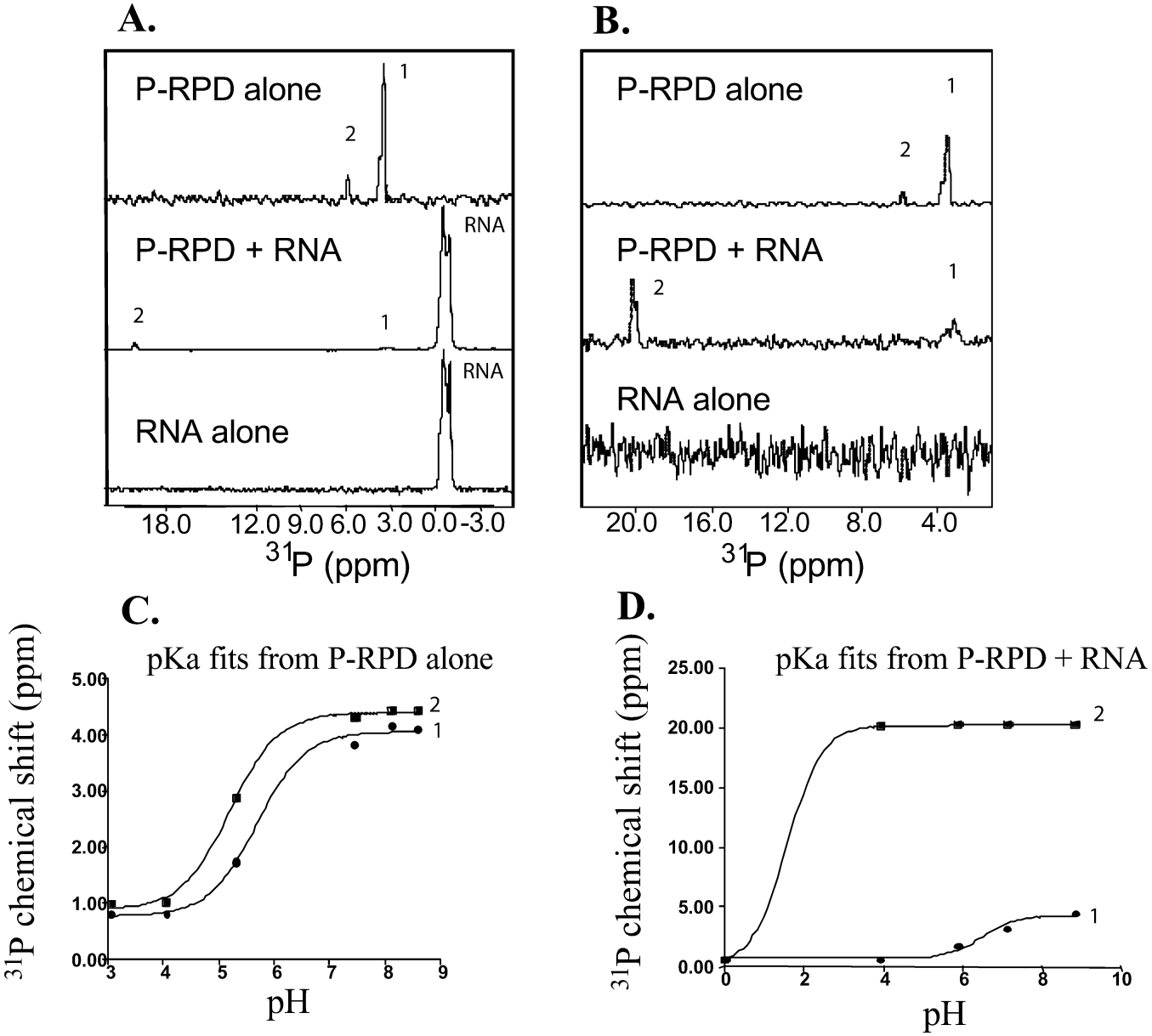

FIGURE 7: Panel A shows the ^{31}P NMR spectra of phosphorylated SLBP RBD in the absence and presence of a 2:1 molar ratio of RNA/protein collected with 20 000 scans at 25 °C on a Varian 500 MHz spectrometer in 20 mM Tris buffer, 50 mM NaCl, 0.1% sodium azide, and 10% ^2H_2O. The ^{31}P spectrum of the RNA alone is shown as a control. In panel B, the region corresponding to the phosphorylated protein resonances is expanded. In panel C, the pK_a values for the P-serine and P-threonine resonances were determined for the free RPD and in panel D for the ^{31}P resonances of the RPD in the RNA complex by monitoring the change in chemical shift as a function of pH. The titration curves for the corresponding resonances in panels A and B are denoted as 1 and 2. The titration curves were fit to the Henderson−Hasselbach equation, and the pK_a values are summarized in Table 2. The different curves for each titration represent different fits to the data to take into account the splitting of the ^{31}P resonance due to proton coupling.

stabilize the RPD as has been observed for the bacteriophage λ N protein (*22*) and the M domain of Ffh (*21*).

Phosphorylation Contributes Favorably to the Overall Stability of the Protein−RNA Complex. Surprisingly, as compared to the unphosphorylated RPD−RNA complex, thermal denaturation experiments show that baculovirus-expressed P-RPD denatures at a T_m that is at least 20 °C higher in the presence of RNA, suggesting that the phosphates stabilize the P-RPD−RNA complex (Figure 3E). pH titrations using ^{31}P NMR are a valuable tool for evaluating the electrostatic contribution of protein phosphorylation to structure and stability. The ^{31}P chemical shift is insensitive to changes in tertiary structure since the phosphate is well shielded by the tetrahedral array of the oxygens but depends on changes in the O−P−O bond angle, as well as the charge on the phosphate (*23*). In the absence of RNA, the ^{31}P NMR spectrum shows two major peaks of different intensities (Figure 7A,B). Both the pK_a values and the observed ^{31}P chemical shifts (Figure 7 and Table 2) are consistent with these resonances representing the ^{31}P-phosphoserine and ^{31}P-phosphothreonine residues in the RPD. The ^{31}P-phosphoprotein resonances can be distinguished from the RNA backbone phosphate resonances since they are shifted downfield relative to the ^{31}P spectrum of RNA alone (Figure 7A,B). Upon titration with RNA, a decrease in peak intensity is observed for resonance 2, and two new peaks appear at ~20 ppm reflecting a change in the protonation state of one or more of these phosphates. The chemical shifts of these

Table 2: ^{31}P Chemical Shifts and pK_a Values[a]

resonance	δ_1	pK_a	δ_2
Random Coil Values (*14*, *15*)			
P-serine	1.16	5.96	5.06/4.7
P-threonine	0.22	6.30	4.51/4.0
P-tyrosine	−3.07	5.96	0.95
P-RPD Alone			
resonance 1	0.79	5.2	4.42
resonance 2	0.99	5.7	4.77
P-RPD + RNA			
resonance 1	0.68	6.28	4.5
resonance 2	20.10	1.56	20.19

[a] Since the relaxation properties of ^{31}P are notoriously affected by CSA and dipolar relaxation mechanism in addition to conformational exchange effects, the peak intensities cannot be integrated to estimate the number of phosphorylated sites. δ_1 and δ_2 represent the chemical shifts of the monoanionic and dianionic phosphate group at low and high pH, respectively. All chemical shifts were referenced to 85% phosphoric acid as an external standard. A positive chemical shift value corresponds to a downfield shift. The pK_a values are calculated as described in Materials and Methods.

downfield-shifted resonances do not change with pH indicating that the phosphates exist in a strong electrostatic interaction with a basic residue such as an arginine or a lysine. In contrast, the chemical shifts of the ^{31}P-serine signals in the free protein titrate as expected with pH, suggesting that this interaction exists only in the RNA-bound form. Hence both the CD data and the ^{31}P titrations provide strong evidence that a favorable electrostatic interaction by at least one phosphoryl group in the SLBP RPD contributes to the stability of the SLBP RPD complex.

The Effect of C-Terminal Phosphorylation Can Be Mimicked by Mutation of Serines to Glutamic Acids. Since the SLBP RPD expressed in *E. coli* is not phosphorylated, it is not possible to evaluate the effect of phosphorylation on the HSQC spectrum. However, as has been shown for the MAP kinases and in the bacterial phosphotransfer protein HPr, the phosphorylation state of a protein can often be mimicked structurally by the side chain carboxylate of an aspartic acid or a glutamic acid (*24*, *25*). We therefore mutated the four serines at the C-terminus to glutamic acids and compared the biophysical and NMR properties of this mutant (4E-RPD) to that of P-RPD. The thermal denaturation profile (Figure 3E) of the free 4E-RPD is similar to that of free RPD in that a broad transition is observed with a midpoint at ∼40 °C. In the presence of stem−loop RNA, the protein melts at a temperature that is between that of the RPD−RNA and P-RPD−RNA complexes suggesting that introduction of four negatively charged moieties at the extreme C-terminus results in increased stabilization of the 4E-RPD as compared to RPD alone. The effect of phosphorylation on stabilization of the structure of the RNA−protein complex is also supported by the differences in the HSQC spectra of 4E-RPD in the presence and absence of RNA. In the absence of RNA, the HSQC spectra are not well resolved (Figure 4B); however, upon the addition of RNA, spectral changes are observed but the HSQC spectra suffer from broad peaks and spectral degeneracy (Figure 4B). However, as expected, the spectral features of the 4E-RPD−RNA complex are much better defined than those observed for the unphosphorylated RPD−RNA (Figure 4A).

In our quest to obtain an NMR-tractable complex, we varied the solution conditions by titrating high concentrations of sodium sulfate into the protein. It has been shown for the protein component of ribonuclease P (*19*) that anions such as sodium sulfate can induce protein folding. The mechanism for anion-induced folding could be through a general electrostatic effect wherein unfavorable charge−charge interactions are overcome in the protein, through the Hofmeister effect, or by binding to specific sites on the protein. Although addition of sodium sulfate does not significantly affect the CD spectrum of SLBP RPD (Figure 3D), we examined the dispersion of resonances in the NMR spectrum as a function of increasing sulfate concentration. As can be seen from Figure 4B, addition of 0.3 M sodium sulfate to 4E-RPD results in folding of the free protein, and uniform cross-peaks are observed for most amides in the HSQC spectrum. A few minor peaks that correspond to the unsulfated form of the protein are also observed in the spectrum. Addition of RNA to this sample results in complete ordering of side chain and backbone resonances (Figure 4B), and NMR spectra have been obtained that are conducive for detailed structure determination. Together these studies show that phosphorylation or addition of high concentrations of sulfate contributes favorably to the stability of the SLBP−RNA complex and both RNA binding and C-terminal phosphorylation are necessary for formation of a stable complex.

DISCUSSION

There is a growing list of proteins, such as FlgM (*26*, *27*), α-synuclein (*28*), the U2AF35 RRM (*29*), and the KIX domain of CREB-binding protein (*30*), that exhibit structural properties similar to those we have described for the SLBP RPD. Evidence from recent NMR studies and biophysical data suggests that many cellular proteins adopt distinct tertiary folds only when complexed to their molecular targets, which can be either another protein or a nucleic acid (RNA or DNA). Many of these proteins have been reported to play important roles in signal transduction and gene expression, and the coupling between folding and binding has been the subject of several recent reviews (*31*−*33*).

However, the majority of known RNA binding proteins are not unfolded but adopt distinct α/β folds. The adaptive binding that is characteristic of RNA binding proteins results either from "ordering" of elements of secondary structure such as loops, α-helices, 3_{10}-helices, or β-ribbons as they are accommodated into the groove of an RNA helix or from the recognition of single-stranded RNA bases by β-sheet surfaces (*34*). Our studies on SLBP show that yet another mode of RNA recognition is possible involving the formation, stabilization, or both of a partly folded protein structure in the presence of a highly ordered RNA hairpin. We have shown here that unlike the bacteriophage λ N protein and the M domain of Ffh, which are completely unfolded, SLBP is natively unfolded in the free state having nativelike secondary structure but not a compact core. The observed properties of dSLBP RPD are not simply due to truncation of the protein since the RPD shows similar spectral features in the context of the full-length protein (Thapar et al., accompanying paper). Strikingly, this structure is protease-resistant.

An unusual feature of SLBP is that formation of a stable dSLBP−RNA complex is coupled to serine phosphorylation.

The only other protein for which the coupling of folding, RNA binding, and phosphorylation has been reported is that of the prokaryotic ribosomal protein L18 (L5 in eukaryotes) where phosphorylation at a serine in the RNA binding domain is necessary for proper folding of the domain in a Mg^{2+}-dependent manner (*35*). The functional homologue of SLBP, polyA-binding protein (PABP), which binds poly(A) tails in the 3′-UTR of all other eukaryotic mRNAs, is also phosphorylated (*36*). However phosphorylation is not important for RNA binding per se, but it promotes cooperative interactions between multiple PABP molecules in the presence of a poly$(A)_{50}$ RNA substrate. Other RNA binding proteins that are known to be regulated by serine phosphorylation include heterogeneous nuclear ribonucleoprotein (hnRNP) K protein (*37*), double-stranded RNA-activated kinase PKR (*38*) and its substrate the hepatitis D virus (HDV), and the coat protein of potato virus A (*39*), among numerous others. In most cases, the role of phosphorylation is to regulate the assembly of multiprotein complexes that are important for RNA processing, translation, export/import, or degradation activities via specific protein−protein interactions and not to regulate RNA binding per se.

Our studies on SLBP also have some parallels with recent reports on CBP KIX−CREB KID and CBP KIX-cMyb interactions (*40*). Although this system is that of a protein−protein complex rather than a protein−RNA complex, it has been shown that in the free state, both the phosphorylated KID and the c-Myb transactivation domains are partly structured and binding to the KIX domain of CBP is coupled to α-helix formation. Phosphorylation of KID in the pKID−KIX system further stabilizes helical structure in the c-Myb transactivation domain and provides intermolecular contacts at the binding interface thereby causing a 20−50-fold increase in binding affinity for the KID domain. In the case of SLBP, the phosphoryl groups in the extreme C-terminus are expected to make intramolecular contacts with basic side chains thereby structuring the RPD.

The biologically relevant question is what is the functional role of C-terminal phosphorylation? Previous in vitro studies on SLBP suggested that although both the RPD and P-RPD are capable of binding RNA, phosphorylation of dSLBP is important for pre-mRNA processing (*10*) and presumably is necessary for binding one of the proteins in the U7 snRNP. The biophysical data reported here suggest that the phosphorylated dSLBP−RNA complex may be more stable, and a longer lived complex may be necessary for the binding of the U7 snRNP multiprotein processing system required for endonucleolytic cleavage of the histone pre-mRNA (*10*). On the other hand, serine phosphorylation may also structure an epitope on the RPD that is specifically recognized by another protein (possibly a U7 snRNP protein) in the processing complex. The kinase that phosphorylates SLBP RPD is unknown, although some of the sites in the sequence DTAKDSNSDSDSD suggest that casein kinase II (CKII) could be a potential kinase (consensus recognition motif for CKII is −X−Ser/Thr−X−X−Asp/Glu−) (*41*).

In summary, our results provide new insights into histone mRNA recognition by dSLBP. The interaction of SLBP with histone mRNA is critical for regulation of histone mRNA metabolism. NMR characterization of the N-terminal domain (in the accompanying paper) and the C-terminal domain suggest that SLBP belongs to the growing family of "intrinsically disordered" proteins and is stabilized only in the presence of another protein or in the presence of an ordered mRNA interface. Phosphorylation further increases the stability of the protein−RNA complex. Future studies will reveal the molecular details important for sequence-specific recognition of histone mRNA by SLBP.

REFERENCES

1. Dominski, Z., and Marzluff, W. F. (1999) Formation of the 3′ end of histone mRNA, *Gene 239*, 1−14.
2. Marzluff, W. F., and Duronio, R. J. (2002) Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences, *Curr. Opin. Cell Biol. 14*, 692−699.
3. Dominski, Z., Erkmann, J. A., Greenland, J. A., and Marzluff, W. F. (2001) Mutations in the RNA binding domain of stem-loop binding protein define separable requirements for RNA binding and for histone pre-mRNA processing, *Mol. Cell. Biol. 21*, 2008−2017.
4. Hanson, R. J., Sun, J., Willis, D. G., and Marzluff, W. F. (1996) Efficient extraction and partial purification of the polyribosome-associated stem−loop binding protein bound to the 3′ end of histone mRNA, *Biochemistry 35*, 2146−2156.
5. Williams, A. S., and Marzluff, W. F. (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.
6. Battle, D. J., and Doudna, J. A. (2001) The stem-loop binding protein forms a highly stable and specific complex with the 3′ stem-loop of histone mRNAs, *RNA 7*, 123−132.
7. DeJong, E. S., Marzluff, W. F., and Nikonowicz, E. P. (2002) NMR structure and dynamics of the RNA-binding site for the histone mRNA stem-loop binding protein, *RNA 8*, 83−96.
8. Zanier, K., Luyten, I., Crombie, C., Muller, B., Schumperli, D., Linge, J. P., Nilges, M., and Sattler, M. (2002) Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression, *RNA 8*, 29−46.
9. Battle, D. J., and Doudna, J. A. (2001) The stem-loop binding protein forms a highly stable and specific complex with the 3′ stem-loop of histone mRNAs, *RNA 7*, 123−132.
10. Dominski, Z., Yang, X. C., Raska, C. S., Santiago, C., Borchers, C. H., Duronio, R. J., and Marzluff, W. F. (2002) 3′ end processing of *Drosophila melanogaster* histone pre-mRNAs: requirement for phosphorylated *Drosophila* stem-loop binding protein and co-evolution of the histone pre-mRNA processing system, *Mol. Cell. Biol. 22*, 6648−6660.
11. Raska, C. S., Parker, C. E., Dominski, Z., Marzluff, W. F., Glish, G. L., Pope, R. M., and Borchers, C. H. (2002) Direct MALDI-MS/MS of phosphopeptides affinity-bound to immobilized metal ion affinity chromatography beads, *Anal. Chem 74*, 3429−3433.
12. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques, *Biochemistry 38*, 16424−16431.
13. Uversky, V. N. (1993) Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule, *Biochemistry 32*, 13288−13298.
14. Vogel, H. J., and Bridger, W. A. (1983) Phosphorus-31 nuclear magnetic resonance pH titration studies of the phosphoproteins tropomyosin and glycogen phosphorylase a, *Can. J. Biochem. Cell Biol. 61*, 363−369.
15. Bienkiewicz, E. A., and Lumb, K. J. (1999) Random-coil chemical shifts of phosphorylated amino acids, *J. Biomol. NMR 15*, 203−206.
16. Wang, Z. F., Whitfield, M. L., Ingledue, T. C., III, Dominski, Z., and Marzluff, W. F. (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.
17. Sullivan, E., Santiago, C., Parker, E. D., Dominski, Z., Yang, X., Lanzotti, D. J., Ingledue, T. C., Marzluff, W. F., and Duronio, R. J. (2001) Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression, *Genes Dev. 15*, 173−187.
18. Martin, F., Schaller, A., Eglite, S., Schumperli, D., and Muller, 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. 16*, 769−778.
19. Henkels, C. H., Kurz, J. C., Fierke, C. A., and Oas, T. G. (2001) Linked folding and anion binding of the *Bacillus subtilis* ribonuclease P protein, *Biochemistry 40*, 2777−2789.
20. Van Gilst, M. R., and von Hippel, P. H. (1997) Assembly of the N-dependent antitermination complex of phage lambda: NusA and RNA bind independently to different unfolded domains of the N protein, *J. Mol. Biol. 274*, 160−173.
21. Zheng, N., and Gierasch, L. M. (1997) Domain interactions in *E. coli* SRP: stabilization of M domain by RNA is required for effective signal sequence modulation of NG domain, *Mol. Cell 1*, 79−87.
22. Van Gilst, M. R., and von Hippel, P. H. (1997) Assembly of the N-dependent antitermination complex of phage lambda: NusA and RNA bind independently to different unfolded domains of the N protein, *J. Mol. Biol. 274*, 160−173.
23. Gorenstein, D. G., Findlay, J. B., Momii, R. K., Luxon, B. A., and Kar, D. (1976) Temperature dependence of the 31P chemical shifts of nucleic acids. A prode of phosphate ester torsional conformations, *Biochemistry 15*, 3796−3803.
24. Thapar, R., Nicholson, E. M., Rajagopal, P., Waygood, E. B., Scholtz, J. M., and Klevit, R. E. (1996) Influence of N-cap mutations on the structure and stability of *Escherichia coli* HPr, *Biochemistry 35*, 11268−11277.
25. Huang, W., Kessler, D. S., and Erikson, R. L. (1995) Biochemical and biological analysis of Mek1 phosphorylation site mutants, *Mol. Biol. Cell 6*, 237−245.
26. Daughdrill, G. W., Hanely, L. J., and Dahlquist, F. W. (1998) The C-terminal half of the anti-sigma factor FlgM contains a dynamic equilibrium solution structure favoring helical conformations, *Biochemistry 37*, 1076−1082.
27. Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T., and Dahlquist, F. W. (1997) The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, sigma 28, *Nat. Struct. Biol. 4*, 285−291.
28. Morar, A. S., Olteanu, A., Young, G. B., and Pielak, G. J. (2001) Solvent-induced collapse of alpha-synuclein and acid-denatured cytochrome *c*, *Protein Sci. 10*, 2195−2199.
29. Kellenberger, E., Stier, G., and Sattler, M. (2002) Induced folding of the U2AF35 RRM upon binding to U2AF65, *FEBS Lett. 528*, 171−176.
30. Radhakrishnan, I., Perez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions, *Cell 91*, 741−752.
31. Dyson, H. J., and Wright, P. E. (1998) Equilibrium NMR studies of unfolded and partially folded proteins, *Nat. Struct. Biol. 5* (Suppl), 499−503.
32. Williamson, J. R. (2000) Induced fit in RNA-protein recognition, *Nat. Struct. Biol. 7*, 834−837.
33. Frankel, A. D., and Smith, C. A. (1998) Induced folding in RNA-protein recognition: more than a simple molecular handshake, *Cell 92*, 149−151.
34. Draper, D. E. (1999) Themes in RNA-protein recognition, *J. Mol. Biol. 293*, 255−270.
35. Bloemink, M. J., and Moore, P. B. (1999) Phosphorylation of ribosomal protein L18 is required for its folding and binding to 5S rRNA, *Biochemistry 38*, 13385−13390.
36. Le, H., Browning, K. S., and Gallie, D. R. (2000) The phosphorylation state of poly(A)-binding protein specifies its binding to poly(A) RNA and its interaction with eukaryotic initiation factor (eIF) 4F, eIFiso4F, and eIF4B, *J. Biol. Chem 275*, 17452−17462.
37. Ostrowski, J., Schullery, D. S., Denisenko, O. N., Higaki, Y., Watts, J., Aebersold, R., Stempka, L., Gschwendt, M., and Bomsztyk, K. (2000) Role of tyrosine phosphorylation in the regulation of the interaction of heterogeneous nuclear ribonucleoprotein K protein with its protein and RNA partners, *J. Biol. Chem 275*, 3619−3628.
38. Chen, C. W., Tsay, Y. G., Wu, H. L., Lee, C. H., Chen, D. S., and Chen, P. J. (2002) The double-stranded RNA-activated kinase, PKR, can phosphorylate hepatitis D virus small delta antigen at functional serine and threonine residues, *J. Biol. Chem. 277*, 33058−33067.
39. Ivanov, K. I., Puustinen, P., Merits, A., Saarma, M., and Makinen, K. (2001) Phosphorylation down-regulates the RNA binding function of the coat protein of potato virus A, *J. Biol. Chem*. *276*, 13530−13540.
40. Zor, T., Mayr, B. M., Dyson, H. J., Montminy, M. R., and Wright, P. E. (2002) Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators, *J. Biol. Chem. 277*, 42241−42248.
41. Peyerin, W., Ackermann, K., and Lorenz, P. (1996) *Biochemistry of Signal Transduction and Regulation*, pp 117−143, VCH Publishers, Inc., New York.

BI036315J