The N-Terminal Domain of the *Drosophila* Histone mRNA Binding Protein, SLBP, Is Intrinsically Disordered with Nascent Helical Structure[†,‡]

Roopa Thapar,[*,§] Geoffrey A. Mueller,[||] and William F. Marzluff[*,§,⊥]

Department of Biochemistry and Biophysics and Department of Biology and Program in Molecular Biology and Biotechnology, University of North Carolina, Chapel Hill, North Carolina 27599, and Laboratory of Structural Biology, National Institute of Environmental Health Sciences, P.O. Box 12233, MD-MR-01, RTP, North Carolina 27709

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

ABSTRACT: Stem−loop binding protein (SLBP) is a 31 kDa protein that is central to the regulation of histone mRNAs and is highly conserved in metazoans. In vertebrates, the N-terminal domain of SLBP has sequence determinants necessary for histone mRNA translation, SLBP degradation, cyclin binding, and histone mRNA import. We have used high-resolution NMR spectroscopy and circular dichroism to characterize the structural and dynamic features of this domain of SLBP from *Drosophila* (dSLBP). We report that the N-terminal domain of dSLBP is stably unfolded but has nascent helical structure at physiological pH and native-like solution conditions. The conformational and dynamic properties of the isolated domain are mimicked in a longer 175-residue region of the N-terminus, as well as in the full-length protein. Complete resonance assignments, secondary structure propensity, and motional properties of a 91-residue N-terminal domain (G17−K108) of dSLBP are reported here. The deviation of $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, and $^{13}C^{\beta}$ chemical shifts from random coil reveals that there are four regions between residues I28−A45, S50−L57, S66−G75, and F91−N96 that have helical propensity. These regions also have small but positive heteronuclear NOEs, interresidue d_{NN} NOEs, and small but significant protection from solvent exchange. However the lack of medium- and long-range NOEs in 3D ^{15}N- and ^{13}C-edited spectra, fast amide proton exchange rates (all greater than 1 s^{-1}), and long ^{15}N relaxation (T_1, T_2) times suggest that the domain from dSLBP does not adopt a well-defined tertiary fold. The backbone residual dipolar couplings (RDCs) for this domain are small and lie close to 0 Hz (±2 Hz) for most residues with no well-defined periodicity. The implications of this unfolded state for the function of dSLBP in regulating histone metabolism are discussed.

The stem−loop binding protein (SLBP),[1] which binds the conserved 26 nucleotide hairpin at the 3′ end of all metazoan replication-dependent histone mRNAs, is important for regulation of histone mRNA processing, translation, and degradation. SLBP has been isolated and cloned from mammals (*1*, *2*), *Drosophila* (*3*), *Caenorhabditis elegans* (*4*, *5*), *Xenopus laevis* (*6*), and *Strongylocentrotus purpuratus* (*7*). In vertebrate and *Drosophila* SLBPs, the region required for RNA binding and pre-mRNA processing is located toward the C-terminal end, whereas the N-terminus of mammalian SLBP and *Xenopus* SLBP1 (xSLBP1) has been shown to be important for translation of histone mRNA (*8*), regulation of SLBP degradation (*9*), cyclin binding (*10*), and nuclear import of histone mRNAs (*11*). In the nucleus, SLBP facilitates histone pre-mRNA processing by stabilizing U7 small nuclear ribonucleoprotein (snRNP) binding to the histone pre-mRNA (*12*). SLBP remains bound to the mature histone mRNA, which is transported to the cytoplasm for translation of histone proteins. Since 70% of the cell-cycle regulation of histone gene expression is modulated post-transcriptionally (*13*) and SLBP is the critical posttranscriptional regulator (*10*), it is important to understand how SLBP carries out its multiple functions.

Figure 1A shows the domain organization of dSLBP identified upon the basis of functional analysis of the protein as well as limited proteolysis of the full-length protein (*15*). dSLBP has a two-domain structure with the RNA binding and processing domain (residues 175−276) in the C-terminal half of the molecule. It has been shown that a 13 amino acid region in the N-terminus of vertebrate SLBPs is required for activation of translation (*8*) (Figure 1B). This sequence is highly conserved in mammals and sea urchins (*6*), and a loose sequence conservation of this motif is found in SLBP

[†] Supported by National Institutes of Health Research Grant GM58961 to W.F.M and a supplement to NIH Grant GM29832 to W.F.M and R.T.

[‡] The HN, ^{15}N, $^{13}C'$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and side chain aliphatic carbon chemical shifts have been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu) under BMRB accession number 5894.

* To whom correspondence should be addressed. E-mail addresses: thapar@email.unc.edu; marzluff@med.unc.edu. Phone: (919) 962-2139. Fax: (919) 962−1274.

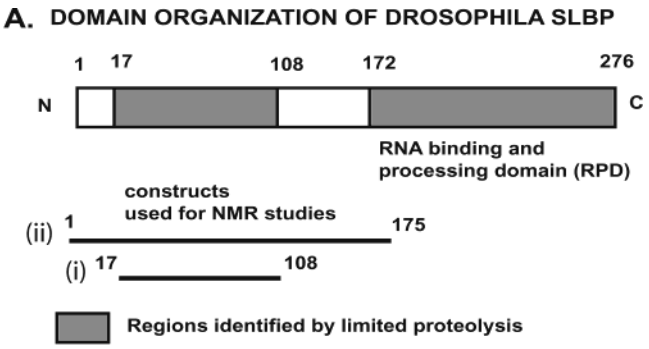
[§] Department of Biochemistry and Biophysics, University of North Carolina.

[||] National Institute of Environmental Health Sciences.

[⊥] Department of Biology and Program in Molecular Biology and Biotechnology, University of North Carolina.

[1] Abbreviations: SLBP, stem−loop binding protein; RPD, RNA binding and processing domain; RBD, RNA binding domain; HSQC, heteronuclear single quantum coherence; 3′ UTR, 3′ untranslated region; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; RDC, residual dipolar couplings; MALDI, matrix-assisted laser desorption ionization; IPAP, in-phase antiphase.

10.1021/bi036314r CCC: $27.50
Published on Web 07/02/2004



B. MOTIF IMPORTANT FOR TRANSLATION INITIATION

dSLBP	^{46}EFGHS-DEAS^{54}
hSLBP	^{73}DWASAVEEDE^{82}
xSLBP1	^{72}DWGSAVEEDE^{81}
ceSLBP	^{61}SWAEVTEEDD^{70}
suSLBP	^{191}DWAVQVEEFE^{200}
cSLBP	^{107}DWFDQVEKSN^{116}

FIGURE 1: Part A depicts the domain structure of dSLBP. The *Drosophila* orthologue of SLBP has a two-domain structure as is delineated in this study. The RNA binding domain is present at the C-terminal end extending from residues 174−261. Residues 261−276 have previously been shown to be important for pre-mRNA processing but not for RNA binding per se (*43*). No function has yet been assigned to residues 1−173. The two constructs used for our studies are (i) residues 17−108, identified by limited proteolysis and mass spectrometry and used for detailed NMR characterization, and (ii) residues 1−175, used for in vivo analysis as well as characterization by CD and NMR. Part B shows the comparison of the amino acid sequence required for translation of histone mRNAs in vertebrate SLBPs with similar regions in other SLBPs. The underlined residues have been shown to be critical determinants for activation of translation (N. Gulseren, R. Thapar, and W. F. Marzluff, unpublished). Alignment of the relevant region in SLBP proteins from human (hSLBP), *Xenopus* (xSLBP1), sea urchin (suSLBP), *Ciona* (cSLBP), *Drosophila* (dSLBP), and *C. elegans* (ceSLBP) is depicted.

from lower eukaryotes, *Drosophila* (*3*) and *C. elegans* (*4*, *5*).

Similar to mammalian SLBPs, evidence from *Drosophila* genetics suggests that dSLBP also shuttles between the nucleus and the cytoplasm implicating a cytoplasmic role for the N-terminal domain of dSLBP (*14*). Moreover *Slbp* mutant embryos expressing a C-terminal fragment of *Slbp* that comprises only the RNA binding and processing domain (RPD) is unable to complement the lethal *Slbp* mutant phenotype (D. J. Lanzotti, Ph.D. thesis, University of North Carolina at Chapel Hill). Although this deletion mutant is capable of restoring histone processing, it is unable to restore viability in vivo suggesting that it is defective in an essential function of dSLBP, most likely its ability to activate translation of histone mRNA.

SLBP fulfills for histone mRNA the regulatory roles in translation and degradation functions that poly(A) binding protein (PABP) fulfills for polyadenylated mRNAs. Therefore to provide further insights into the function of the N-terminus of dSLBP and as part of our ongoing effort to understand structure−function relationships involving dSLBP, we have characterized the conformational properties of this domain using circular dichroism and NMR spectroscopy. We show that the dSLBP N-terminus (residues 1−175) is not stably folded in solution. Detailed NMR characterization of the structural and motional properties of a 91-residue "structural domain" reveals four regions in this domain that prefer helical conformations in solution. Therefore dSLBP belongs to the growing family of intrinsically disordered or unfolded proteins that undergo a folding transition upon recognition of their biological targets.

MATERIALS AND METHODS

Protein Expression and Purification and Sample Preparation. To identify stable domains of dSLBP, we subjected full-length, wild-type, baculovirus-expressed SLBP to limited proteolysis and MALDI mass spectrometry as described (*15*). The residues corresponding to the N-terminal domain (NT, residues G17−K108) identified by limited proteolysis and residues M1−K175 were subcloned into the *Nde*1 and *Bam*H1 restriction sites of the vector pET15b. Full-length *Drosophila* SLBP-4E (wherein the last four serines in the extreme C-terminus were mutated to glutamic acids (*13*)) was subcloned into the *Nde*1/*Xho*1 sites of the bacterial expression vector pET28a. All proteins were expressed with either N- or C-terminal His tags, which does not influence the function of SLBP. Unlabeled and uniformly ^{15}N-labeled full-length dSLBP-4E, dSLBP NT, and dSLBP NT175 were expressed by growing BL21(DE3) RIL CodonPlus cells (Stratagene) transformed with the respective pET vectors at 20 °C in the presence of 1 mM IPTG in Luria broth or minimal media. The resulting full-length and N-terminal proteins were expressed with an N-terminal hexahistidine tag with the sequence MGSSHHHHHHSSGLVPRGSH, which has no apparent effect on the structure of the protein. A single doubly labeled (^{15}N, ^{13}C) sample of the N-terminal domain (residues 17−108) was prepared by growing *Escherichia coli* BL21(DE3) RIL CodonPlus cells (Stratagene) transformed with the plasmid in minimal medium containing 1.0 g/L of 99% $^{15}NH_4Cl$ and 3 g/L of 99% ^{13}C-glucose as the sole sources of nitrogen and carbon, respectively. The protein was further purified using immobilized metal-affinity chromatography using Ni^{2+} ions followed by gel filtration over an S100 column. In addition, a specifically ^{15}N-labeled G/S sample, and a uniformly ^{15}N-labeled sample with selectively unlabeled lysine were used for assignment verification. The proteins were >95% pure as judged by SDS−PAGE. The identity of all constructs was confirmed by DNA sequencing as well as electrospray mass spectrometery. All triple resonance experiments were performed on a single 2 mM uniformly (^{15}N,^{13}C)-labeled NMR sample in 50 mM Tris buffer, 50 mM NaCl, 5 mM DTT, 0.1% sodium azide, pH 6.0, in $H_2O/^2H_2O$ (90:10).

NMR Measurements. NMR data was collected at 25 °C and pH 7.0 on Varian Inova 600 MHz spectrometers equipped with a 5 mm z-gradient triple resonance probe. Heteronuclear NMR experiments used for obtaining backbone and side chain assignments of the N-terminal domain (residues 17−108) of SLBP consisted of 2D (^{15}N,^{1}H) HSQC and 3D HNCA, HN(CO)CA, HNCO, HNCACB, CC(CO)NH total correlation spectroscopy (TOCSY), HC(CO)NH TOCSY (τ_m = 21 ms), and HCCH TOCSY (isotropic mixing time of 18 ms). Triple resonance experiments were collected as previously described (*16*−*19*). Spectra were processed using the program Felix 98.2 (Accelrys). The proton chemical shift was referenced using DSS as an internal standard, whereas the ^{15}N and ^{13}C

chemical shifts were referenced using the respective gyromagnetic ratios. NOE information was obtained from a 3D (^{15}N,^{1}H) NOESY−HSQC (mixing time of 75 ms) and a 3D (^{13}C, ^{1}H) NOESY−HSQC (mixing time of 100 ms) and $^{3}J_{NH\alpha}$ coupling information from a 3D HNHA. The chemical shift indices were calculated using the program CSI (*20*).

We tried to measure hydrogen exchange rates of slowly exchanging amides at 5 °C and pH* 7.0 for both NT (17−108) and NT175 proteins by ^{1}H−^{2}H exchange using short 15 min HSQCs, but none of the amides were protected 2 min after exchange into $^{2}H_2O$. The measurement of exchange rates of rapidly exchanging amide protons ($k_{ex} > 1$ s^{-1}) was carried out by the method of Spera, Ikura, and Bax (*21*). Two identical (^{1}H,^{15}N) HSQC spectra were recorded either with presaturation of water between scans or using gradients for water suppression. Pairs of spectra were collected at 25 °C at pH values of 6.01, 6.90, and 7.74 ± 0.05. Hydrogen exchange rates (k_{obs}) were calculated at pH 7.0 as outlined in Spera et al. (*21*). The intrinsic hydrogen exchange rates (k_{int}) were calculated by the method of Bai et al. (*22*) using an in-house Fortran program written for this purpose. Protection factors (*P*) for individual amides are defined as the ratio k_{int}/k_{obs}.

^{15}N T_1 and T_2 relaxation rates and steady-state {^{1}H}−^{15}N NOEs were measured using the pulse sequences described by Farrow et al. (*23*). All relaxation rates and steady-state {^{1}H}−^{15}N NOE experiments were measured at 25 °C at 600 MHz. For the T_1 experiments, 512 × 256 complex points were collected with relaxation delay times of 5, 50, 150, 250, 350, 400, 450, 500, 750, and 1000 ms, and the T_2 experiments were performed with relaxation delay times of 14.4, 28.8, 57.6, 86.4, 100.8, 115.2, 129.6, 144, 158.4, and 172.8 ms. The steady-state {^{1}H}−^{15}N NOE experiment was performed by collecting two identical datasets recorded with and without proton presaturation. The saturation period consisting of a train of 120° proton pulses was 4.9 s. The interscan delay period was 5 s. All spectra were recorded with 256 complex t_1 increments in the indirect dimension. The spectral widths were 4801 Hz in the amide proton and 1100 Hz in the ^{15}N dimension. For the T_1 experiments, 180° proton pulses were applied every 5 ms during the relaxation delay period, and for the T_2 experiments, 180° pulses were applied every 625 ms and proton 180° pulses every 10 ms. Peak intensities were extracted from all spectra, and the data were fit using the program Curvefit (Palmer, A., 1998).

Residual dipolar couplings (RDC) were measured by preparing a 2 mM (^{15}N,^{13}C)-labeled sample of the shorter 91 amino acid domain (SLBP NT) in the presence of Pf1 phage (Asla). The same NMR sample was used for the isotropic and anisotropic measurements. About 10 mg of Pf1 phage was used in the anisotropic sample as verified from the 10 Hz splitting of the HDO signal. For the measurement of $^{1}J_{NH}$ couplings (^{1}H,^{15}N) IPAP−HSQC NMR spectra were recorded at 600 MHz. The spectra were acquired with spectral widths of 5000 Hz in ^{1}H and 1100 Hz in ^{15}N. RDCs were determined from the difference in one bond $^{1}J_{NH}$ couplings between isotropic and anisotropic samples (RDC = $^{1}J_{NH}$(aligned) − $^{1}J_{NH}$(unaligned)). Data were processed and analyzed using NMRPipe. The periodicity in the $^{1}D_{NH}$ couplings was assessed according to Mesleh and Opella (*24*).

Circular Dichroism. CD scans were performed in 20 mM sodium phosphate buffer, pH 7.0, at 25 °C. The conformational stability of the N-terminal domains 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.

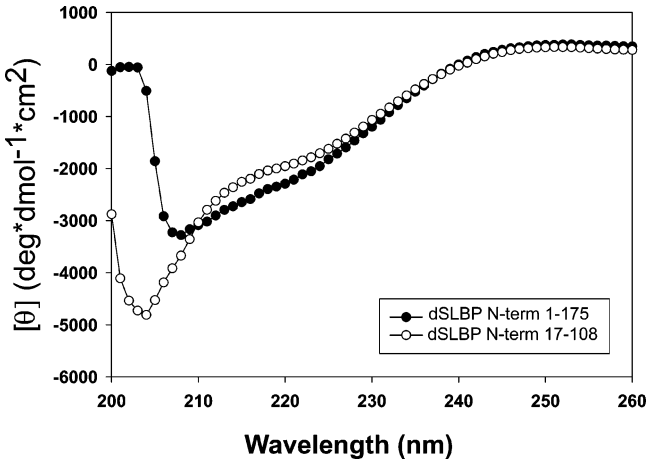


FIGURE 2: Far UV CD scans of SLBP-NT (residues 17−108) and SLBP-NT (residues 1−175). Both samples of SLBP-NT were in 20 mM sodium phosphate buffer, pH 7.0, at 50 μM. The CD scans were performed at 25 °C.

RESULTS

Drosophila SLBP (dSLBP) gives two protease-resistant fragments when digested with a number of proteases (*15*). One fragment corresponds to the C-terminus of dSLBP, which contains the region necessary for histone pre-mRNA processing consisting of residues H172−D276, and the second fragment contains a portion of the amino terminus from residues G17−K108 (Figure 1). The remainder of the protein (amino acids 108−172) is very glycine- and serine-rich and likely unstructured.

Circular Dichroism Suggests Helical Secondary Structure Exists for the N-Terminus of dSLBP.

Figure 2 shows a CD spectrum of the two dSLBP N-terminal constructs that we have employed for our studies measured at neutral pH and 25 °C. The CD spectrum for the longer N-terminal construct (1−175) shows double minima at 209 and 222 nm, which is indicative of helical structure. The shorter N-terminal domain shows minima at 204 and 222 nm suggesting that some helical structure is present; however, the minimum at 204 nm is shifted from the expected value of 209 nm reflecting contributions from random coil. The mean molar ellipticity at 222 nm is a rough measure of helical content in proteins, and this value is quite small for both N-terminal constructs ($\theta_{222} = -2110$ and $\theta_{222} = -1842$ for residues 1−175 and 17−108, respectively). Therefore, only a small percentage of the molecule exists in a helical conformation at neutral pH and 25 °C. Surprisingly, no appreciable increase in helical content (as interpreted from the ellipticity at 222 nm) is observed for the 175-residue construct compared to the 91 residue domain suggesting that residues that lie outside the 91-residue domain are mostly flexible and unstructured. This is in agreement with our results from limited proteolysis experiments that showed residues 17−108 were resistant to proteolytic degradation and with NMR characterization of the longer 175-residue

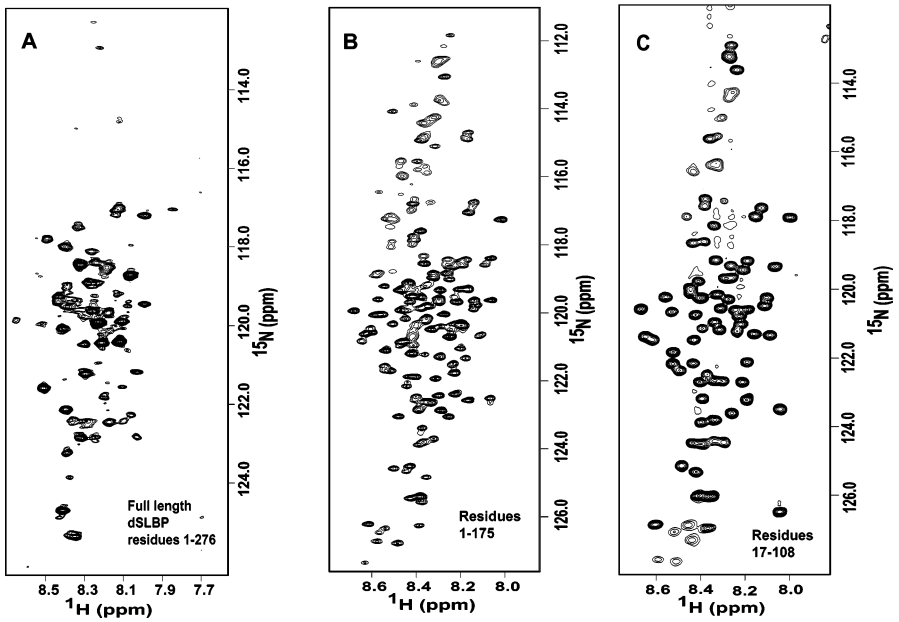


FIGURE 3: (^{1}H,^{15}N) HSQC spectra collected at a frequency of 600 MHz at 25 °C and pH 6.0 of (A) full-length SLBP-4E mutant (residues 1−276), (B) the N-terminal domain (residues 1−175), and (C) the shorter (91-residue) N-terminal domain. The ^{15}N and ^{1}H spectral widths were set to 1100 and 4801 Hz, respectively, to maximize resolution. All spectra were collected under identical spectral conditions.

fragment reported here. The observations from the CD scans were further emphasized in the CD-monitored thermal denaturation profiles of the two constructs (data not shown) that did not show a cooperative unfolding transition as would be expected if the domains were well ordered.

NMR Characterization of a Smaller 91-Residue Fragment of dSLBP NT

Comparison of (^{1}H,^{15}N) HSQC Spectra of Full-Length and N-Terminal Fragments of dSLBP. Figure 3 shows a comparison of the (^{1}H,^{15}N) HSQC spectra of full-length dSLBP, residues M1−N175 of the N-terminal domain, and residues G17−K108 of the N-terminus. Remarkably, the chemical shifts of resonances in both the full-length protein and the 175-residue and 91-residue N-terminal domains have a very narrow dispersion in the amide proton region of 0.4 ppm that is centered around the random coil chemical shift (between 8.2 and 8.4 ppm) (Figure 3). The (^{1}H,^{15}N)-HSQC spectrum of the full-length SLBP protein reveals that only 50/262 (276 residues − 14 prolines) of the expected cross-peaks are observed in the spectrum at 600 MHz, and these cross-peaks are also centered between 8.2 and 8.4 ppm. The remaining peaks are absent from the spectrum either due to chemical exchange broadening especially of the C-terminal residues (from the RNA binding and processing domain or RPD) or due to rapid amide exchange. We show in the accompanying paper (*15*) that the RNA binding domain is only partly folded and the cross-peaks for residues in this domain are exchange-broadened in HSQC spectra. The RPD is fully functional in RNA binding. Therefore the properties of both the N-terminal and C-terminal domains are reflected in the full-length protein.

Backbone Chemical Shifts. Since the domain that we identified from limited proteolysis experiments corresponded to residues G17−K108 and there is no appreciable change in the CD spectrum between the 91-residue and 175-residue constructs, we decided to focus our NMR studies on the smaller domain. Moreover, many of the cross-peaks for this domain could be superimposed onto the spectrum of the longer 175-residue domain suggesting that any residual structure is likely to be similar between the two N-terminal constructs. To characterize the structural features of dSLBP-NT, we completely assigned the backbone and side chain aliphatic resonances of this domain using standard double and triple resonance techniques (3D HNCA, HNCOCA, HNCACB, CBCA(CO)NH, HNCO, CC(CO)NH TOCSY, HCC(CO)NH TOCSY, and HCCH TOCSY) as described in Materials and Methods. The assigned spectrum of the smaller 91-residue domain used for this study is shown in Figure 4. Despite the narrow chemical shift dispersion in the amide proton dimension of 0.4 ppm, the (^{1}H,^{15}N) HSQC spectrum of the 91-residue protein collected with a sweep width of 4.8 kHz in the amide proton region and 1.1 kHz in the ^{15}N dimension resulted in very good resolution of all cross-peaks. Triple resonance experiments that utilized the resolution in the ^{15}N and ^{13}C′ dimensions to resolve potential spectral overlap in $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ enabled complete assignment of all observable backbone and side chain resonances for the smaller N-terminal fragment (G17−K108) (Figure 4). Residues that remain unassigned include residues S23−S27, N58−S60, S63, and C64 that are not observed in the NMR spectra. The assignment of glycine and serine residues was also confirmed from specifically labeled samples. Amide cross-peaks corresponding to nine glycines have been aliased in the spectrum and hence appear downfield shifted in Figure 4.

The chemical shifts of backbone $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and ^{13}C′ nuclei are strongly correlated with secondary structure and are commonly used to predict backbone dihedral angles in a folded protein. For an unfolded protein, the deviation of $^{1}H^{\alpha}$, $^{13}C^{\alpha}$, $^{13}C^{\beta}$, and ^{13}C′ chemical shifts from random coil are the best measure of secondary structure propensity since they are sensitive to small deviations from random coil values. A downfield shift of the $^{13}C^{\alpha}$ resonance by ~2.5 ppm

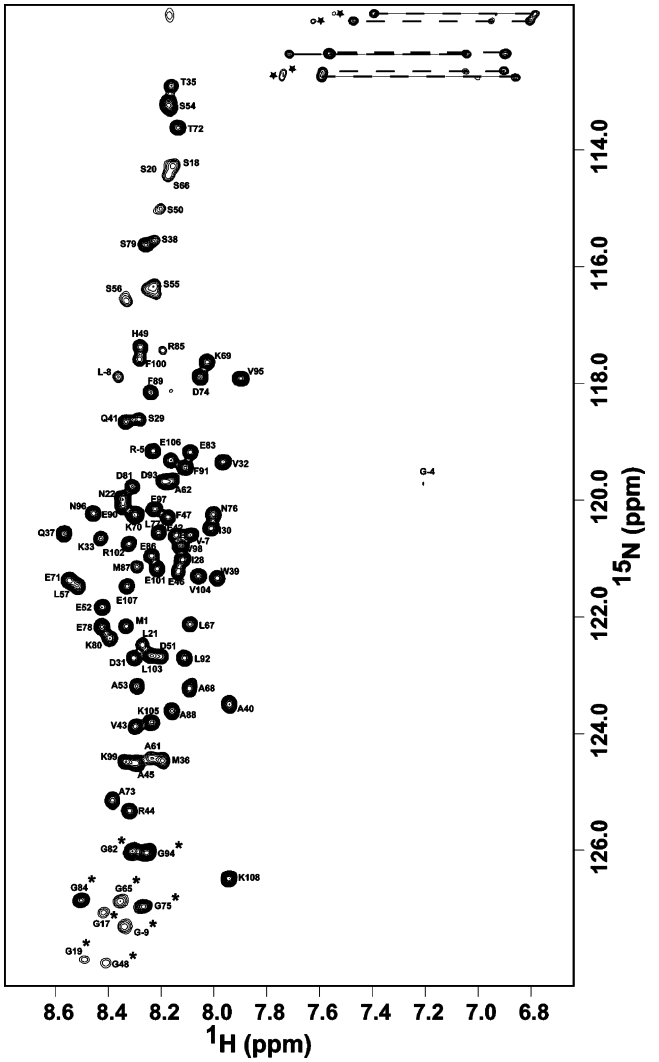
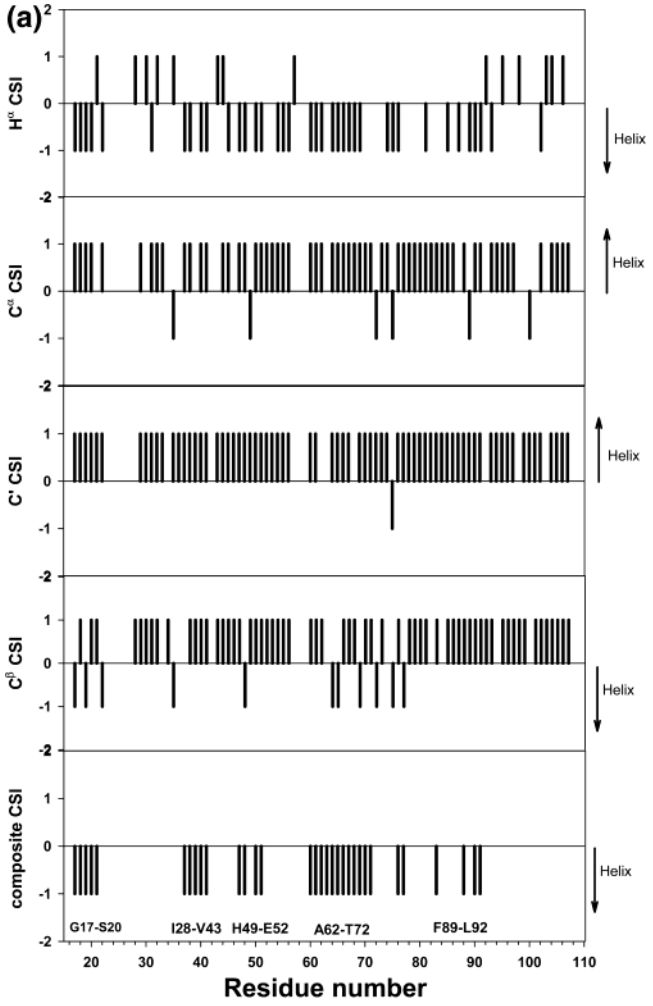


FIGURE 4: (^{1}H,^{15}N)-HSQC spectrum of uniformly ^{15}N-labeled sample of the N-terminal domain (residues 17−108) of SLBP collected at a ^{1}H frequency of 600 MHz at pH 6.0 and 25 °C. This is the minimal N-terminal domain that was identified using limited proteolysis and mass spectrometry. The spectrum was collected with a 2 mM sample with 128 increments and eight scans per t_1 increment (total acquisition time is 1 h, 8 min). The residue-specific assignments of cross-peaks arising from the backbone amide groups are indicated. Resonances marked with negative numbers denote nonnative residues from the vector that was left as a solubility enhancement tag. Glycine residues that have been aliased in the spectrum are denoted by asterisks (∗). Side chain amide protons of asparagine and glutamine residues are shown connected by dashed lines. Four additional side chain resonances of weak intensity are observed in the spectrum, and these are denoted as ★.



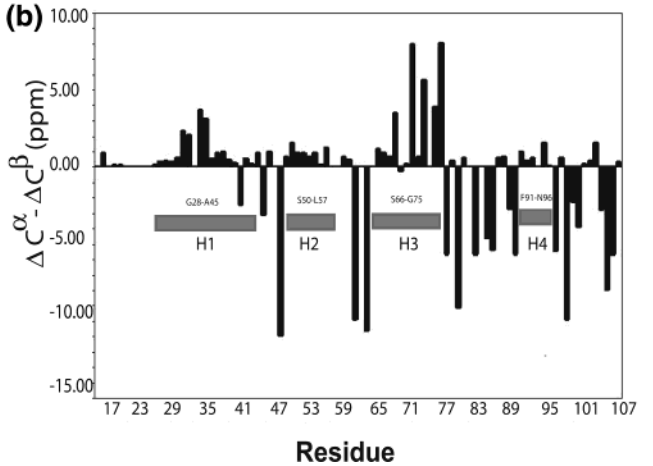


FIGURE 5: In part A, the deviation of H^α , $^3C^\alpha$, $^3C^\beta$, and $^3C'$ chemical shift values from random coil sequence-corrected as previously described is depicted as the corresponding chemical shift indices (CSI) calculated using the program CSI. In part B, the ($\Delta^{13}C^\alpha - \Delta^{13}C^\beta$) chemical shift difference measured from a 3D HNCACB spectrum is shown as a function of residue number. H1 comprises residues G28−A45; H2 comprises residues S50−L57; H3 comprises residues S66−G75; H4 comprises residues F91−N96.

is indicative of α-helical structure, whereas an upfield shift by ∼1.7 ppm is suggestive of β-sheet. Figure 5A shows the chemical shift indices (CSI) for H^α, C^α, C^β, and C′ and the composite CSI for all heteronuclei, which indicate that there are five putative helical regions in the molecule that correspond to G17−S20, I28−V43, H49−E52, A62−T72, and F89−L92. The correlation of $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts with secondary structure further increases by calculating the ($\Delta^{13}C^\alpha - \Delta^{13}C^\beta$) shift, where the Δ shift corresponds to the deviation from random coil (*25*, *26*). Figure 5B shows the ($\Delta^{13}C^\alpha - \Delta^{13}C^\beta$) shift values for residues in SLBP-NT. There are four regions that show positive values of ($\Delta^{13}C^\alpha - \Delta^{13}C^\beta$) shifts, and these correspond to residues I28−A45, S50−L57, S66−G75, and F91−N96. No clear trend is observed for the remaining residues, suggesting that these are present in an extended conformation or a coil. The magnitude of the ($\Delta^{13}C^\alpha - \Delta^{13}C^\beta$) shift is largest for residues S66−L77 with the largest difference being 8 ppm for E71,

Table 1: Vicinal $^3J_{NH\alpha}$ Coupling Constants and $^1H-^{15}N$ Residual Dipolar Couplings ($^1D_{NH}$) (in Hz) for Residues G17−K108 of dSLBP

residue	$^3J_{NH\alpha}$	D_{NH}	*a*	residue	$^3J_{NH\alpha}$	D_{NH}	*a*	residue	$^3J_{NH\alpha}$	D_{NH}	*a*
G17		*b*		E52	7.0	0	H	M87		+2.4	
S18		+0.6		A53	6.0	+1.8	H	A88	6.6	−1.2	
G19		*b*		S54	6.6	−1.2	H	F89	7.7	−1.8	
S20	7.5	*b*		S55		−2.4	H	E90		+1.2	
L21		−1.2		S56		−1.2	H	F91		−1.8	H
N22		*b*		L57	6.4	+1.2	H	L92	8.6	+2.4	H
S23		*b*		N58		*b*		D93		+0.6	H
S24		*b*		S59		*b*		G94	7.8	0	H
A25		*b*		S60		*b*		V95	7.8	+1.8	H
S26		*b*		A61	7.8	0		N96	7.2	−1.8	H
S27		*b*		A62		+1.8		E97	5.5	0	
I28	7.0	+1.8	H	S63		*b*		V98	7.1	+1.2	
S29	7.0	+3.0	H	C64		*b*		K99	6.8	+3.0	
I30	7.3	+1.2	H	G65		*b*		F100	7.5	+1.8	
D31	6.8	0	H	S66		−2.4		E101	7.2	−3.0	
V32	8.1	0	H	L67	5.8	+1.2	H	R102	7.6	+2.4	
K33	6.3	+3.0	H	A68	5.7	+1.2	H	L103	8.2	0	
P34		*b*	H	K69	8.0	0	H	V104	8.1	0.6	
T35		+1.1	H	K70	6.5	−2.4	H	K105	7.9	+2.4	
M36		+4.9	H	E71	6.6	+1.2	H	E106	7.8	−1.2	
Q37	6.5	+1.2	H	T72	8.4	−1.2	H	E107	7.8	+2.4	
S38		N. D.	H	A73	7.3	−1.8	H	K108	8.3	0	
W39	7.9	+1.2	H	D74	6.3	+1.8	H				
A40	7.2	−1.2	H	G75	7.8	0					
Q41	6.4	−0.6	H	N76	7.2	0					
E42		−3.6	H	L77	6.2	0					
V43	6.7	−1.8	H	E78	7.8	−1.2					
R44	7.2	+1.8	H	S79	6.5	+1.2					
A45	7.6	+1.8	H	K80	8.7	−3.0					
E46	7.0	0		D81	6.2	−1.8					
F47	6.5	0		G82	5.9	+1.2					
G48		*b*		E83	7.1	+1.2					
H49	8.2	+1.2		G84	7.9	−1.2					
S50		0	H	R85		*b*					
D51		+1.2	H	E86		−1.2					

[a] Refers to helical secondary structure propensity as is estimated from all of the NMR parameters combined. [b] Not determined.

whereas all other residues have values <3 ppm, which suggests conformational averaging. Therefore the pattern of the CSI and ($\Delta^{13}C^{\alpha} - \Delta^{13}C^{\beta}$) shifts is consistent with helix being present for four regions in the molecule; however, the small magnitude of the shifts suggests conformational averaging exists for these helical regions as well.

Coupling Constants and NOEs. The $^3J_{NH\alpha}$ vicinal coupling constant is known to be empirically correlated with the backbone ϕ angle via the Karplus relationship. Residues with $^3J_{NH\alpha}$ < 6 Hz are correlated with helical conformation, whereas $^3J_{NH\alpha}$ values of >8 Hz are correlated with β-strands. Residues with $^3J_{NH\alpha}$ values between 6 and 8 Hz are expected to either lack stable secondary structure or have unusual dihedral angles associated with these couplings. $^3J_{NH\alpha}$ couplings were measured for 58 residues in a 3D HNHA spectrum for which the cross-peaks were well resolved and the peaks were not exchange-broadened. Of these, only three residues have $^3J_{NH\alpha}$ values (Table 1) of <6 Hz suggestive of helical or turn structure. These residues are L67, A68, and E97. The remaining residues have couplings between 6 and 8 Hz indicating that it is likely that the backbone dihedral ϕ for these residues exchanges rapidly between an ensemble of conformations in solution. The presence of helical structure can be identified from the pattern of d_{NN}, $d_{\alpha N}$, $d_{\beta N}$, and $d_{\alpha\beta}$ short and $d_{\alpha_i N_{(i+3)}}$, $d_{\alpha_i \beta_{(i+3)}}$ medium-range NOEs in NOESY spectra. A 75 ms 3D ^{15}N-edited NOESY spectrum

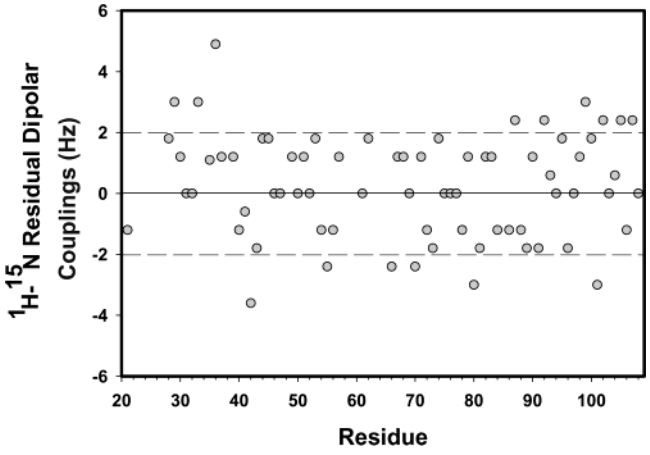


FIGURE 6: Residual dipolar couplings measured for the N-terminal domain of *Drosophila* SLBP. Small but significant changes in dipolar couplings are observed in the sequence suggesting differences in the relative alignment of residues in the protein. The experimentally measured dipolar couplings are plotted as a function of residue number.

and a ^{13}C-edited NOESY spectrum collected on this molecule show that weak d_{NN}'s exist for some residues, but no medium-range NOEs were identified. In addition, the spectra show a paucity of long-range NOEs, and the intraresidue correlations are strong suggesting that the molecule is largely unfolded. Figure 9 summarizes the limited number of d_{NN} NOEs observed in the four putative helical regions in the molecule.

$^1D_{NH}$ Residual Dipolar Couplings. Residual dipolar couplings (RDCs) are important probes of backbone structure and dynamics (*27*, *28*). In well-folded molecules, RDCs provide orientation information for individual N−H bond vectors with respect to the magnetic field such that in an ideal helix, all the NH bond vectors align similarly along the helical axis and hence would be oriented with similar angular coordinates with respect to the principal axis of the alignment tensor. This results in distinct sinusoidal patterns termed "dipolar waves" where the periodicity would be 3.6 for ideal α-helical geometry. To test whether the RDCs suggest the presence of either stable or nascent helical structure, SLBP-NT was aligned in the presence of ∼10 mg of Pf1 phage under similar conditions used for the analysis of other NMR parameters (pH 6.0 and 50 mM NaCl at 25 °C). The final concentration of phage in the NMR sample was assessed by the splitting of the HDO signal, which was 10 Hz. Small shifts in the spectrum of dSLBP-NT caused by anisotropic averaging of the chemical shift tensor confirmed the weak alignment of dSLBP-NT. The chemical shifts of amide cross-peaks did not change dramatically, and no significant broadening of resonances was observed indicating that the SLBP-NT domain does not associate strongly with the phage. Figure 6 and Table 1 summarize the RDC values for backbone N−H vectors extracted from IPAP−HSQC spectra. The RDC values for SLBP-NT are small and range from −2.4 to +4.9 Hz. Most of the RDC values lie within ±2 Hz, and there is no apparent periodicity that is characteristic of helices. When the data is fit to the "dipolar wave" equations as described by Mesleh et al. (*24*), most of the protein has a very low "periodicity" score (<5), indicating that no one stretch of residues is readily distinguishable as helical compared to any other region. Since the

uncertainity in the measurements is ~1 Hz, we cannot ascertain whether the small RDC values reflect the lack of persistent long-range structure in SLBP-NT or averaging about the N−H bond. Since the other NMR parameters suggest that SLBP-NT lacks stable tertiary structure and the backbone is highly flexible, the latter is probably correct, although we cannot rule out the former.

Shortle and colleagues observed large $^1D_{NH}$ values for Δ131Δ staphylococcal nuclease in 8 M urea for the Δ131Δ staphylococcal nuclease mutant in the presence of increasing amounts of urea (*44*). In this study, the authors suggested that the RDC values reflected persistent structure in the protein and the $^1D_{NH}$ values were not affected by local disorder. This result is contrary to initial expectations that isotropic averaging about the NH bond vector, which is characteristic of unfolded proteins should result in RDC values that lie close to zero. It is also contrary to recent experimental results on a themally denatured GB1 domain that found no significant RDCs. However, a recent theoretical study suggests that neglecting end effects, RDCs will be small, negative, and monotonic for random flight chains of length similar to SLBP-NT (*29*). It is tempting to conclude therefore that the variability in the measurements presented here reflect some persistent structure in SLBP-NT, although the error in the measurements is large compared to the range of couplings.

Amide Exchange Rates. The protection of backbone amides from exchange with solvent is observed either when the peptide is structured in the vicinity of the exchanging NH and lies in a hydrogen bond or when the amide is buried in a hydrophobic core and hence inaccessible to solvent. Amide exchange methods have long been used to detect local structure as well as serve as reporter groups in protein folding. We performed H−D exchange experiments using NMR for the 91 amino acid SLBP-NT and the longer 175 amino acid domain. The experiment was performed at 5 °C to decrease k^{HX} at a pH of 7.0 and with a dead time of 2 min. However under these conditions none of the amides are protected from exchange with D_2O suggesting that the amides exchange rapidly on the NMR time scale (the half-life is on the order of milliseconds). We conclude therefore that none of the amides exist in stable, persistent secondary structure. To measure exchange rates of fast-exchanging amides (i.e., where k^{HX} occurs faster than 1 s^{-1}), we utilized the method of Spera, Ikura, and Bax (*21*) whereby the exchange rate can be extracted from the ratios of the cross-peak intensities in spectra collected with and without water saturation. Figure 7A shows portions of the (^1H,^{15}N) HSQC spectra of SLBP-NT collected at different pH values in the presence of either gradients or presaturation for water suppression. As is evident from Figure 7A, the cross-peak intensities decrease considerably when the pH is increased from 6.01 to 7.74. At neutral pH (pH 6.9), there is also a large contribution to the amide proton exchange rate due to cross-saturation by water. By measuring spectra at different pH values, the true k^{HX} can be calculated as described in Materials and Methods. To take into account the inductive and steric contribution of the side chain at the ($i - 1$) position to the exchange rate, protection factors (defined as the ratio k^{HX}/k^{int}) were calculated for all amides. Figure 7B shows the logarithms of the protection factors plotted against residue number. The strongest protection from exchange is observed

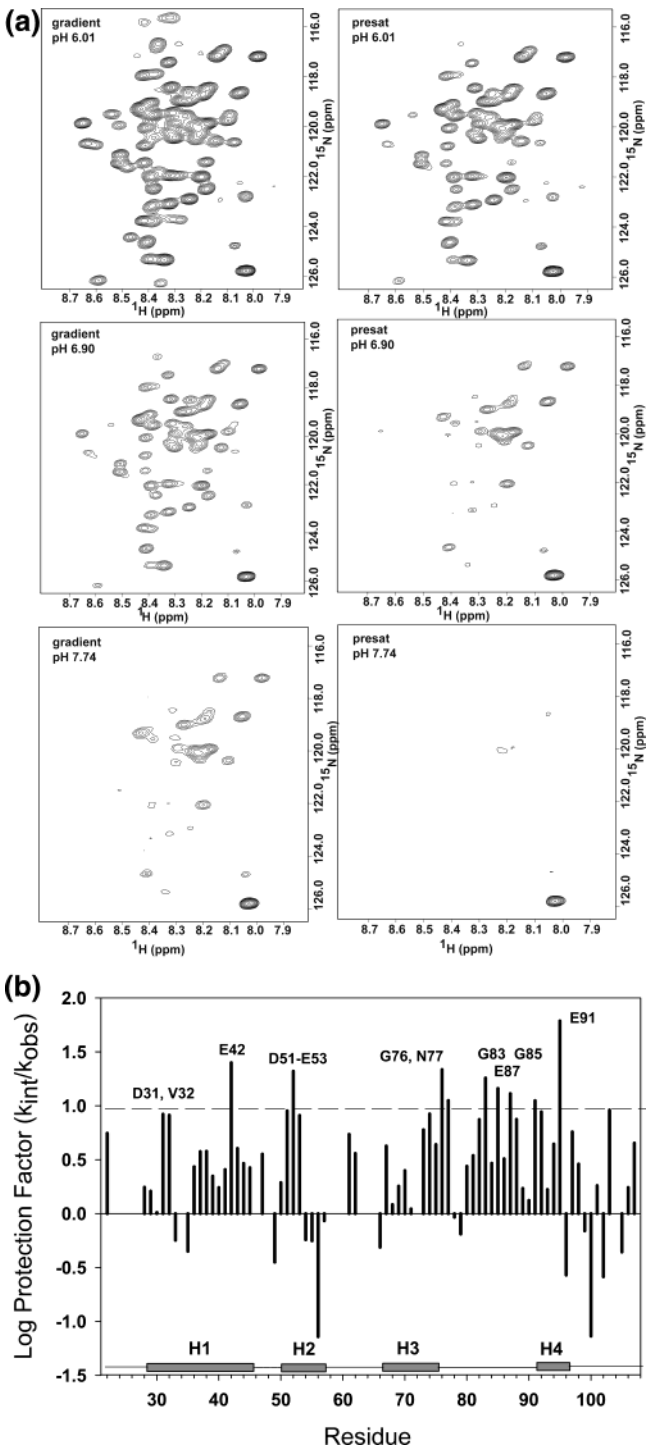


FIGURE 7: Part A shows measurement of fast-exchanging backbone amides for residues F17−K108 from HSQC spectra collected with and without presaturation of water at three different pH values. The left panels show HSQC spectra collected without presaturation of H_2O, and the right panels show spectra collected with presaturation during the relaxation delay period. Part B shows amide proton protection factors determined for all amides in the N-terminal domain (residues 17−108) of dSLBP.

for residues that lie in regions of helical propensity as is predicted from heteronuclear chemical shifts. However even for these residues, only a 10-fold protection from exchange relative to the intrinsic exchange rate for a random coil peptide is observed (as compared to protection by 10^3−10^6-fold for folded proteins), reflecting the highly dynamic behavior of the backbone of this domain. Neverthe-

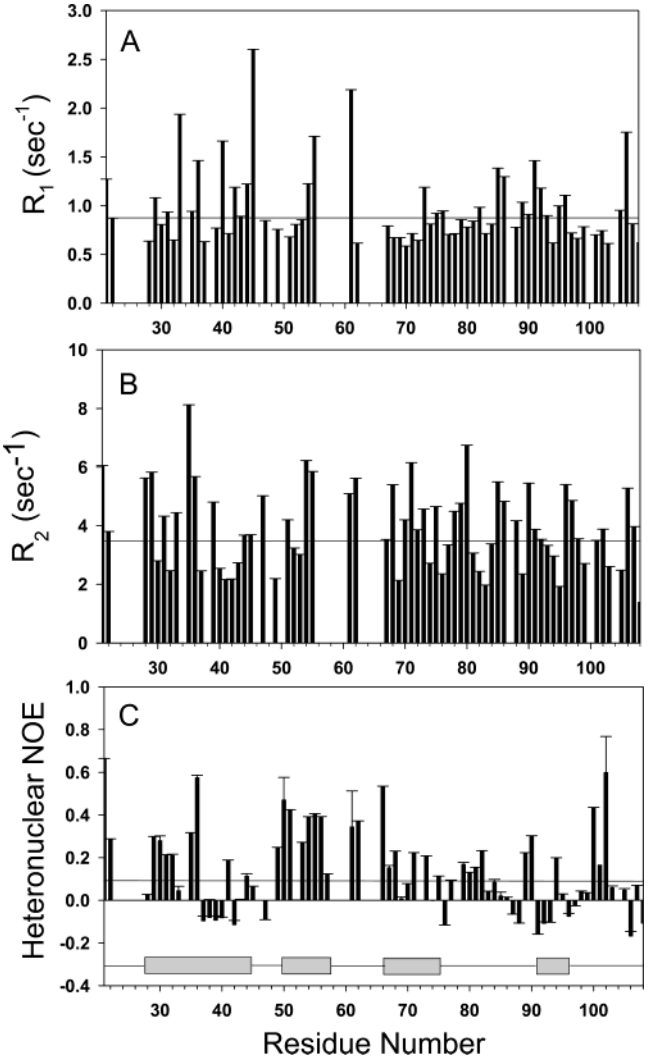


FIGURE 8: Backbone dynamics for the N-terminal domain of SLBP. Plot of backbone ^{15}N R_1 (A) and R_2 (B) and {^{1}H} − ^{15}N heteronuclear NOE (C) values for the N-terminal domain of dSLBP determined at 600 MHz.

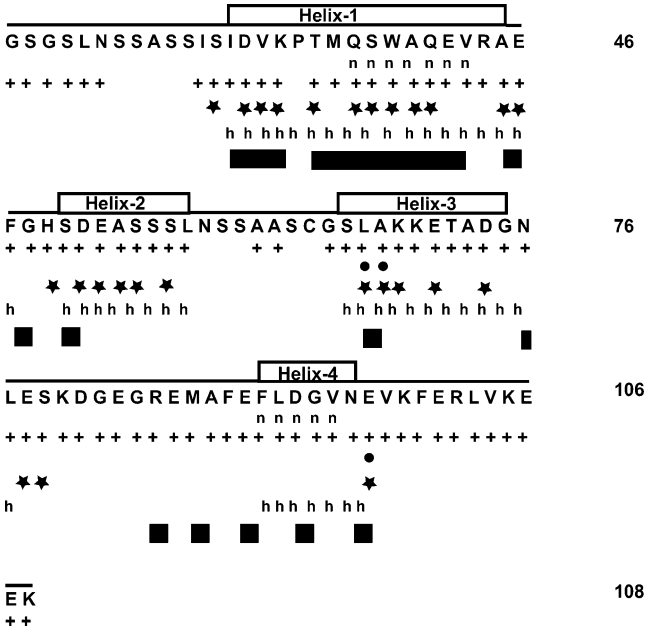


FIGURE 9: Summary of residual secondary structure and amino acid sequence of the N-terminal domain of *Drosophila* SLBP. Residues that have been assigned are designated by a plus sign; the pattern of d_{NN}'s is denoted by closed rectangles; residues with $^3J_{NH\alpha} < 6$ Hz are shown with filled circles; residues predicted to be helical in the chemical shift index are denoted as h for helix. Residues showing more than 10-fold protection from exchange relative to random coil are depicted as star; those with higher than average heteronuclear NOEs are depicted as n. On the basis of all the NMR parameters combined, the following regions are expected to have some residual helical structure: I28−A45, S50−L57, S66−G75, and F91−N96.

less it does suggest that transient structure exists for these helical regions.

^{15}N Relaxation Rates and Heteronuclear NOEs. The motional properties of the N-terminal domain of dSLBP on the nanosecond to picosecond time scale were studied by measuring the ^{15}N T_1 and T_2 relaxation rates for 67/91 backbone amides. Cross-peaks that were of weak intensity or in regions of spectral overlap were left out of the analysis. ^{15}N−^{1}H heteronuclear NOEs and ^{15}N T_1 and T_2 relaxation rates were determined from intensity-modulated spectra as described in Materials and Methods and are shown in Figure 8. The T_1 (Figure 8A) and T_2 (Figure 8B) relaxation times were much longer than those expected for a small globular domain of ∼10 kDa. The ^{15}N T_1 ranged from 384 to 1704 ms with an average value measured over 67 amides of 1151.2 ± 40.11 ms, whereas the ^{15}N T_2 ranged from 123 to 719 ms with a mean value of 287.3 ± 13.84 ms. The ^{15}N−^{1}H heteronuclear NOE, which is a good indicator of internal motion in a protein backbone was negative for most backbone amides (Figure 8C) and ranged from −0.17 to +0.66. The mean ^{15}N−^{1}H heteronuclear NOE value was small and was determined to be 0.14 ± 0.023. The relaxation parameters are also consistent with the presence of residual structure in the four "helical" regions in the protein. Some residues in the regions I28−Q41 (helix-1) and S66−G75 (helix-2) show small but positive heteronuclear NOE values, and these values are also smaller than average for residues S50−S56 and S66−L77. The differences in backbone motion between the helical regions and the rest of the molecule were also apparent in the relaxation times (Figure 8A,B), which were lower than the average T_1 and T_2 values determined for the entire molecule. The values of the relaxation parameters determined clearly indicate that rapid internal motion exists for all amides, consistent with the protein being mostly unfolded. These data are also consistent with the NMR structural parameters that suggest that partly folded residual helical propensity exists in four regions of this domain.

NMR Characterization of the Longer N-Terminal Domain Comprising Residues 1−175. One possible explanation for the observed properties of the smaller 91 amino acid domain is that it is an artifact due to premature truncation of the protein. Therefore we also analyzed NMR data for the longer 175 residue domain. As shown in Figure 2, the HSQC spectrum of the 175-residue domain also has a narrow dispersion of chemical shifts in the amide proton region, and many of the resonances from the smaller domain superimpose onto the spectra of the 175-mer. Measurement of H/D exchange rates for the longer 175-residue protein at 5 °C and pH 7.0 shows that similar to the smaller domain, all amides exchange rapidly with solvent (with an exchange rate of >0.01 s^{-1}, which is the lower limit of exchange in this experiment) and none of the amides are protected from

exchange 2 min after the addition of D_2O. It is therefore unlikely that stable secondary or tertiary structure exists for residues 1−175. In addition, as part of another study, we have also measured equilibrium $^2H/^1H$ isotope effects for the 175 residue protein (data not shown). The equilibrium fractionation factor (ϕ) or enrichment of protium vs deuterium in a protein at equilibrium is correlated with hydrogen bond strength such that amides in strong H-bonds accumulate protium ($\phi < 1$), whereas amides in weak H-bonds accumulate deuterium ($\phi > 1$). In general, α-helical residues have lower ϕ values than amides in β-sheet residues. The results indicate that 86% of all amides in the 175-residue protein have ϕ values of >1 with values greater than the average ϕ value observed in proteins with well-defined structure. Therefore both the H/D exchange rates and the $^2H/^1H$ isotope effects suggest that amide hydrogen bonds in residues 1−175 of the N-terminal domain of dSLBP are weak and the protein is not stably folded under native-like solution conditions.

DISCUSSION

We have characterized the structural and dynamic properties of the N-terminal domain of dSLBP. Our data shows that dSLBP is an intrinsically unfolded protein and its N-terminal domain has transient helical structure in four regions of the molecule. The term "intrinsically unfolded" or "intrinsic disorder" has been used to describe polypeptide chains that are by their inherent nature not structured as opposed to polypeptides that are normally structured but can be induced to unfold using denaturants. The spectroscopic properties of the N-terminal domain of SLBP are very similar to that of FlgM and β-dystroglycan, both of which have been well characterized by NMR and are regarded as being intrinsically or natively unfolded (*30*). The motional properties of the N-terminal domain of SLBP are very different from that of the C-terminal RPD, which is also not stably folded but of which the HSQC spectrum is suggestive of slower exchange between multiple conformations (*15*). The N-terminal domain of SLBP gives one cross-peak per residue indicating that, although unfolded, the different conformations are in fast exchange on the NMR time scale. The exchange broadening observed for the C-terminal RPD suggests that the time scale of exchange is much slower, that is, occurs on the millisecond time scale, and the RPD has features of a molten globule-like state (*15*). The features of both the N-terminal domain and the C-terminal RPD are present in the full-length protein; therefore, our results on the individual domains are relevant to the properties of the intact molecule.

Biological Relevance of Intrinsic Disorder. Bioinformatics approaches estimate that 35−51% of proteins in the eukaryotic genome have regions of low complexity or global disorder in which 40 or more consecutive residues are predicted to be unfolded (*31*). Frequently these regions are deemed to be important for recognition of other proteins or nucleic acids although there are very few examples wherein dynamics has been directly correlated with biological function (*32*). Global disorder in intrinsically unfolded proteins is less well characterized than local disorder and the role of global disorder in modulating protein function is not well understood. Therefore structural characterization of these proteins is necessary to understand the importance of disorder in regulating biological function. It has been suggested that intrinsically unfolded proteins such as dSLBP may control affinity and specificity separately during a binding event such that they pay an energetic penalty in conformational entropy but are able to maintain highly specific interactions with their biological targets (*33*, *34*). It has also been proposed that the ability to maintain an unfolded state allows these proteins/domains to adopt unfavorable backbone conformations in the bound state that would be energetically very costly if the protein were folded into a unique structure. More structural and biophysical studies are therefore needed on free and bound forms of unfolded proteins to fully understand the biological relevance of being unfolded. Our characterization of the "free" state of SLBP reported herein helps to understand the conformational transitions that the protein must undergo when it binds eukaryotic translation initiation factors.

SLBP is an adapter-like molecule that has been proposed to bind other proteins as well as RNA (*12*). In the cell, there is a pool of SLBP in the nucleus that is used to process histone pre-mRNA as it is transcribed. When SLBP binds to the pre-mRNA, it can then interact with a protein in the U7 snRNP, assisting in the recruitment of U7 snRNP to the pre-mRNA (*35*). Once the histone mRNA is processed SLBP remains bound to the 3′ end of the mRNA and accompanies the RNA to the cytoplasm. The N-terminal region of SLBP must interact transiently with a number of different proteins to stimulate translation, is a substrate for multiple protein kinases, is required for nuclear import of SLBP, and is required for targeting SLBP for degradation. Therefore a functional advantage of being unfolded may be the ability of dSLBP to bind numerous targets (*36*). SLBP is also a phosphoprotein and is phosphorylated at numerous sites in the N- and C-terminal domains (*37*, *43*). Phosphorylation of the N-terminal domain of human SLBP has been shown to be important for degradation of hSLBP. *Drosophila* SLBP has been shown to be phosphorylated at T120 in the N-terminus; however, the functional significance of this phosphorylation is unclear (*14*). Since phosphorylation sites are frequently disordered in proteins, this may be another reason dSLBP evolved to be inherently unfolded.

Comparison of NMR Parameters of dSLBP NT with NMR Studies of Other Unfolded Molecules. From a biophysical stand-point, dSLBP resembles a folding intermediate where transient secondary structure persists due to local interactions between residues, but long-range interactions are not present (*38*−*40*; for reviews, see refs *41* and *42*). Our data on dSLBP are similar to those reported for the native state of FlgM in which the C-terminal domain is unfolded but favors helical conformations (*33*). Similar to FlgM, some of the residues have helical propensity as is estimated from heteronuclear chemical shifts; however, the relaxation times T_1 and T_2 are long suggesting that a compact globular state does not exist. The native state of dSLBP also has properties similar to the denatured state of the Δ131Δ construct of SNase in that the amide exchange rates are fast suggesting that hydrogen bonds are not stable and there is little stable packing of residues on the nanosecond to picosecond time scale as is estimated from relaxation data (*39*). However, in contrast to Δ131Δ staphylococcal nuclease, we do not observe large RDC values that are indicative of persistent long-range structure, and our results are more consistent with recent studies on a thermally

denatured domain of GB1 that showed no significant $^1D_{NH}$ couplings (*45*).

NMR spectroscopy is uniquely suited for the study of unfolded proteins; however, the number of NMR parameters available that are sensitive to the detection of residual structure is small. We have characterized the N-terminal domain of SLBP in detail. Our results show that deviation of heteronuclear chemical shifts from random coil remains the single most important measure of residual structure since the $^3J_{NH\alpha}$ and $^1D_{NH}$ couplings are usually averaged and few medium and long-range NOEs are observed. We also show that the backbone RDC values in an unfolded molecule such as SLBP lie close to zero hertz. It is important to characterize residual structure in these unfolded proteins, both from the viewpoint of protein folding and for understanding disorder−function relationships as opposed to structure−function relationships. The challenge in the coming years will be to develop methods that allow better characterization of the residual and persistent structure in these molecules.

In summary, we have provided the first insight into the molecular features of dSLBP. Future structural characterization of dSLBP bound to protein and RNA targets will help to understand the structural changes that occur in dSLBP upon binding translation initiation factors and the role of intrinsic disorder in molecular recognition for this biologically important protein.

ACKNOWLEDGMENT

We thank Michael Mesleh for sharing the MatLab scripts for analysis of the $^1D_{NH}$ data.

Note Added after ASAP Posting. The paper was originally posted on July 2, 2004. Figure 3 was replaced, and the article was reposted on July 6, 2004.

REFERENCES

1. 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.
2. 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.
3. 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.
4. Pettitt, J., Crombie, C., Schumperli, D., and Muller, B. (2002) The *Caenorhabditis elegans* histone hairpin-binding protein is required for core histone expression and is essential for embryonic and postembryonic cell division, *J. Cell Sci. 115*, 857−866.
5. Kodama, Y., Rothman, J. H., Sugimoto, A., and Yamamoto, M. (2002) The stem-loop binding protein CDL-1 is required for chromosome condensation, progression of cell death and morphogenesis in *Caenorhabditis elegans*, *Development 129*, 187−196.
6. Wang, Z. F., Ingledue, T. C., Dominski, Z., Sanchez, R., and Marzluff, W. F. (1999) Two Xenopus proteins that bind the 3′ end of histone mRNA: implications for translational control of histone synthesis during oogenesis, *Mol. Cell. Biol. 19*, 835−845.
7. Robertson, A. J., Howard, J. T., Dominski, Z., Schnackenberg, B. J., Sumerel, J. L., McCarthy, J. J., Coffman, J. A., and Marzluff, W. F. (2004) The sea urchin stem-loop-binding protein: a maternally expressed protein that probably functions in expression of multiple classes of histone mRNA, *Nucleic Acids Res. 32*, 811−818.
8. Sanchez, R., and Marzluff, W. F. (2002) The stem-loop binding protein is required for efficient translation of histone mRNA in vivo and in vitro, *Mol. Cell. Biol. 22*, 7093−7104.
9. Zheng, L., Dominski, Z., Yang, X. C., Elms, P., Raska, C. S., Borchers, C. H., and Marzluff, W. F. (2003) Phosphorylation of stem-loop binding protein (SLBP) on two threonines triggers degradation of SLBP, the sole cell cycle-regulated factor required for regulation of histone mRNA processing, at the end of S phase, *Mol. Cell. Biol. 23*, 1590−1601.
10. Whitfield, M. L., Zheng, L. X., Baldwin, A., Ohta, T., Hurt, M. M., and Marzluff, W. F. (2000) Stem-loop binding protein, the protein that binds the 3′ end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms, *Mol. Cell. Biol. 20*, 4188−4198.
11. Erkmann, J. A. (2003) Ph.D. Thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC.
12. 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.
13. Harris, M. E., Bohni, R., Schneiderman, M. H., Ramamurthy, L., Schumperli, D., and Marzluff, W. F. (1991) Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps, *Mol. Cell. Biol. 11*, 2416−2424.
14. Lanzotti, D. J., Kupsco, J. M., Yang, X. C., Dominski, Z., Marzluff, W. F., and Duronio, R. J. (2004) Drosophila stem-loop binding protein intracellular localization is mediated by phosphorylation and is required for cell cycle-regulated histone mRNA expression, *Mol. Biol. Cell 15*, 1112−1123.
15. Thapar, R., Marzluff, W. F., and Redinbo, M. R. (2004) Electrostatic contribution of serine phosphorylation to the *Drosophila* SLBP−histone mRNA complex, *Biochemistry 43*, 9401−9412.
16. Grzesiek, S., Dobeli, H., Gentz, R., Garotta, G., Labhardt, A. M., and Bax, A. (1992) ^1H, ^{13}C, and ^{15}N NMR backbone assignments and secondary structure of human interferon-γ, *Biochemistry 31*, 8180−8190.
17. Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins, *J. Biomol. NMR 3*, 185−204.
18. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-Enhanced Triple-Resonance 3-Dimensional NMR Experiments with Improved Sensitivity, *J. Magn. Reson. Ser. B 103*, 203−216.
19. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments, *J. Biomol. NMR 3*, 225−231.
20. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry 31*, 1647−1651.
21. Spera, S., Ikura, M., and Bax, A. (1991) Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment, *J. Biomol. NMR 1*, 155−165.
22. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75−86.
23. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation, *Biochemistry 33*, 5984−6003.
24. Mesleh, M. F., and Opella, S. J. (2003) Dipolar Waves as NMR maps of helices in proteins, *J. Magn. Reson. 163*, 288−299.
25. Constantine, K. L., Mueller, L., Goldfarb, V., Wittekind, M., Metzler, W. J., Yanchunas, J., Jr., Robertson, J. G., Malley, M. F., Friedrichs, M. S., and Farmer, B. T. (1997) Characterization of NADP+ binding to perdeuterated MurB: backbone atom NMR assignments and chemical-shift changes, *J. Mol. Biol. 267*, 1223−1246.
26. Metzler, W. J., Constantine, K. L., Friedrichs, M. S., Bell, A. J., Ernst, E. G., Lavoie, T. B., and Mueller, L. (1993) Characterization of the three-dimensional solution structure of human profilin: 1H, 13C, and 15N NMR assignments and global folding pattern, *Biochemistry 32*, 13818−13829.
27. Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science 278*, 1111−1114.

28. Prestegard, J. H. (1998) New techniques in structural NMR- - anisotropic interactions, *Nat. Struct. Biol. 5* (Suppl), 517−522.
29. Louhivuori, M., Paakkonen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. (2003) On the origin of residual dipolar couplings from denatured proteins, *J. Am. Chem. Soc. 125*, 15647−15650.
30. Bozzi, M., Bianchi, M., Sciandra, F., Paci, M., Giardina, B., Brancaccio, A., and Cicero, D. O. (2003) Structural characterization by NMR of the natively unfolded extracellular domain of beta-dystroglycan: toward the identification of the binding epitope for alpha-dystroglycan, *Biochemistry 42*, 13717−13724.
31. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function, *Biochemistry 41*, 6573−6582.
32. Eisenmesser, E. Z., Bosco, D. A., Akke, M., and Kern, D. (2002) Enzyme dynamics during catalysis, *Science 295*, 1520−1523.
33. 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.
34. Spolar, R. S., and Record, M. T., Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA, *Science 263*, 777−784.
35. Dominski, Z., Erkmann, J. A., Yang, X., Sanchez, R., and Marzluff, W. F. (2002) A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3′-end processing, *Genes Dev. 16*, 58−71.
36. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure−function paradigm, *J. Mol. Biol. 293*, 321−331.
37. 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.
38. Dill, K. A., and Shortle, D. (1991) Denatured states of proteins, *Annu. Rev. Biochem. 60*, 795−825.
39. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures, *J. Mol. Biol. 268*, 170−184.
40. Choy, W. Y., and Forman-Kay, J. D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain, *J. Mol. Biol. 308*, 1011−1032.
41. Dyson, H. J., and Wright, P. E. (2001) Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states, *Methods Enzymol. 339*, 258−270.
42. Shortle, D. R. (1996) Structural analysis of non-native states of proteins by NMR methods, *Curr. Opin. Struct. Biol. 6*, 24−30.
43. 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.
44. Shortle, D., and Ackerman, M. S. (2001) Persistence of native-like topology in a denatured protein in 8M urea, *Science 293*, 487−489.
45. Ding, K., Louis, J. M., and Gronenborn, A. M. (2004) Insights into the conformation and dynamics of protein GB1 during folding and unfolding by NMR, *J. Mol. Biol. 335* (5), 1299−1307.

BI036314R