Functional Implications for a Prototypical K-Turn Binding Protein from Structural and Dynamical Studies of 15.5K[†,‡]

Sarah E. Soss[§] and Peter F. Flynn*

Department of Chemistry, University of Utah, Salt Lake City, Utah 84112

Received June 26, 2007; Revised Manuscript Received October 17, 2007

ABSTRACT: The kink-turn (K-turn) motif is recognized and bound by a family of proteins that act as nucleation factors for ribonucleoparticle assembly. The binding of various proteins to a conserved RNA structural motif known as the K-turn has been shown to be an important component of regulation in the ribosome, in the spliceosome, and in RNA modification. 15.5K is a prototypical example of a K-turn binding protein, which has been shown to bind the 5′-U4 stem-loop of the spliceosome and the box C/D motif. We describe the solution NMR structure of free 15.5K, as well as studies of conformational flexibility from ^{15}N NMR relaxation and H/D exchange experiments. The protein appears well-structured aside from conformational fluctuation in α3. Flexibility in fast time scale motions and the observation of limited intermediate and slow motions further characterize the free protein and may suggest local contributions to recognition and binding.

The processing of nascent RNA transcripts is essential to their utility. This processing includes base modifications (methylation and pseudouridinylation) as well as more complex sequence modifications such as splicing. The spliceosome is a dynamic macromolecular complex which undergoes significant rearrangements during the execution of its function (*1*−*3*). The human 15.5kD/NHPX protein, or simply 15.5K, specifically binds to U4 snRNA and is required for assembly of the full U4/U6 snRNP (*4*). 15.5K is highly conserved across species, and many aspects of its function have been studied in the yeast homologue, Snu13p (*5*−*8*). The 15.5K protein binds specifically to the 5'stem-loop of U4 with a K_d of 20 nM, and disruption of this protein−RNA interaction has been shown to block splicing (*4*). In the standard model of U4/U6 snRNP assembly and rearrangement, 15.5K first binds to U4 snRNA, which leads to the subsequent recruitment of the 61K protein and the 20/60/90K protein complex to the U4/U6 RNP (*9*). During the transformation of the spliceosome into a catalytically active state, the base pairs between U4 and U6 snRNA are disrupted and the U4 snRNP is released, creating the U6/U2/pre-mRNA complex which constitutes the catalytic core (*9*).

Other types of RNA processing, such as methylation and pseudouridinylation, are regulated by small nucleolar ribonucleoparticles (snoRNPs) (*10*, *11*). Remarkably, in addition to its function in RNA splicing the 15.5K protein also binds the box B/C and box C/D RNA motifs (*7*, *12*). Similar to its role in splicing, 15.5K takes part in assembly of snoRNPs. These particles function in the nucleolus; therefore, 15.5K is a crucial multifunctional protein that participates in at least two distinct RNA processing events involving multiple RNA targets.

The RNA targets for 15.5K share a novel structural motif known as the kink-turn or K-turn. The K-turn motif appears to be ubiquitous, and examples of this organizational structure have been found in snRNA, snoRNA, the large and small subunits of the ribosome, and even pre-mRNA (*13*). Although these motifs are found throughout the cell, 15.5K and eukaryotic homologues have thus far been found to bind targets exclusively in snRNAs and snoRNAs, whereas the archaeal protein homologue, L7Ae, binds targets in the ribosome in addition to snoRNAs. There is currently no high-resolution structural information for a free K-turn RNA in solution, and based on first principle considerations and direct experimental evidence it is unlikely that the uncomplexed K-turn structure would be stable in solution (*14*).

Previous work on the 15.5K−U4SL complex has shed light on the manner in which 15.5K binds RNA. Hydroxy radical footprinting studies of U4SL RNA conducted by Lührmann and co-workers show a distinct pattern of protected residues, e.g., 29−33 and 43−47 are protected by 15.5K binding (*4*). Structural information about this system comes from work on the homologues Snu13p (*15*) and L7Ae (*16*−*19*), as well as the K-turn binding L30 proteins (*20*−*22*). More importantly, a crystal structure of the human 15.5K has been solved in complex with a 22-nucleotide RNA containing the 5′ stem-loop of U4 snRNA (*23*). Salient features of the complex include an α−β−α sandwich motif for the protein with the RNA-binding interface on one face and an 120° angle between the two base-paired stems of the hairpin RNA. The structure of the 15.5K−RNA complex

[†] This work was supported by a Seed Grant from the University of Utah and the Biological Chemistry Training Grant NIH GM8537 (SES). The Utah Health Sciences NMR facility is supported by grants RR13030 (NIH) and DBI-0002806 (NSF).

[‡] Coordinates have been deposited in the Protein Data Bank under accession code 2JNB.

* Corresponding author. Tel: (801) 581-3828 Fax: (801) 581-8433, E-mail: pfflynn@chem.utah.edu.

[§] Present address: Department of Biochemistry and Center for Structural Biology, Vanderbilt University School of Medicine, Nashville, TN 37232.

10.1021/bi701254q CCC: $37.00
Published on Web 11/29/2007

forms a solid basis for studies involving interactions between K-turn-forming RNAs and the proteins that bind them.

Comparative studies of free 15.5K, free U4SL, and the 15.5K−RNA complex are essential for further progress in elucidation of the role of 15.5K in the regulation of splicing and RNA methylation functions. It has become increasingly clear that induced fit has a central role in the formation of protein−nucleic acid complexes (*24*, *25*), and only a portion of its function in the 15.5K−U4SL complex has been elucidated. The structural work reported thus far does not include a complete free and bound pair for a single protein species, and no complete dynamical studies for a K-turn binding protein have been reported. We demonstrate here that a solution NMR-based approach provides important new insights into the structure, function, and dynamics of 15.5K in the K-turn protein−RNA interaction.

EXPERIMENTAL PROCEDURES

Sample Purification. The amino acid sequence for human 15.5K was back-translated to produce the DNA sequence, employing codon usage optimized for expression in *E. coli*. An N-terminal 6×-His-tag was coengineered into the nucleotide sequence, which was then generated using PCR and cloned into pET11a. This plasmid was used to transform BL21 (DE3) cells, which were then employed for the expression of 15.5K. Cells were grown at 37 °C in M9 minimal media, labeled using $^{15}NH_4Cl$ and/or ^{13}C-glucose, to an optical density of 0.6−0.8 (A_{600}) and then induced with 1 mM IPTG for 4 h to reach a terminal OD of 1.2−1.5. Cells were harvested by centrifugation and stored at −80 °C. After thawing, cells were lysed in Ni-binding buffer (300 mM sodium chloride, 50 mM sodium phosphate, 10 mM imidazole, pH 8.0) containing lysozyme and 1 mM PMSF. Following sonication, the cell extract was treated with RNase A (10 μg/mL) on ice for 20 min. The extract was then cleared by centrifugation and passed through a 0.45 μm filter. The extract was loaded onto a 5-mL Ni-IDA column, which was washed with binding buffer containing 100 mM imidazole, and the protein was then eluted with 500 mM imidazole. The sample was concentrated and separated with a gel filtration column to further purify His-15.5K from the cell extract. The fractions containing the protein were then concentrated and the buffer exchanged into 100 mM sodium chloride, 100 mM sodium phosphate, pH 6.0, using an Amicon stirred cell. This process regularly yields 10−15 mg of purified protein per liter of culture.

NMR Spectroscopy. Samples containing ∼8% D_2O and ∼1.0 mM protein were used for all experiments. Distance restraints were obtained from 3D ^{15}N- and ^{13}C-resolved NOESY-HSQC experiments recorded at 18.5 °C and 800 MHz (^1H). All NMR data were processed using Felix 97 (Accelrys, San Diego CA) and peak lists generated from Sparky (*26*).

All NMR relaxation experiments were recorded at 20 °C on a 1.5 mM sample of His-15.5K in 100 mM sodium chloride, 100 mM sodium phosphate, pH 6.0. Longitudinal (T_1) and transverse (T_2) relaxation of ^{15}N nuclei were measured at 11.7 and 14.0 T, corresponding to ^{15}N Larmor frequencies of 50.68 and 60.78 MHz, using pulse sequences based on the ^{15}N-HSQC experiment (*27*). The $\{^1H\}-^{15}N$ NOE was also measured at both fields, based on ^{15}N-HSQC experiments recorded with and without broadband ^1H saturation for 5 s (*27*). Longitudinal NMR relaxation experiments employed delay times of 45, 85, 130, 206, 286, 376, 477, 582, 798, 998, and 1250 ms at 50.68 MHz and 45, 85, 130, 206, 286, 376, 477, 582, 798, 998, and 1495 ms at 60.78 MHz. Transverse NMR relaxation experiments at 50.68 MHz employed CPMG delays of 8, 16, 24, 32, 48, 64, 80, 104, 136, and 168 ms. The corresponding experiments at 60.78 MHz employed delays of 8.1, 16.2, 24.3, 32.4, 48.6, 64.8, 81, 105, 138, and 170 ms.

Structure Calculation. Structure calculations including automated NOE assignment were completed using CYANA 2.1 (*28*, *29*). Previously determined chemical shift assignments (*30*) were transferred to peaks in the NOESY data before creating the input chemical shift list for calculations. Additional restraints used include torsion angle restraints from TALOS (*31*) and 78 long-range NOEs for which unambiguous assignment was known. The final set of structures was further refined against experimental restraints using CNS (*32*), validated using Pro-check NMR (*33*), and visualized using PyMOL (DeLano Scientific). The entire protein sequence including the N-terminal tag was used in structure calculation, but only the 15.5K protein (residues 1−128) is included in the final structure. The rmsd of the structure was calculated based on residues 9−127, constituting the ordered portion of the protein. The structure and selected associated data have been deposited to the Protein Data Bank as 2JNB.

^{15}N Relaxation Data Analysis. Resonances were quantified using their intensities, and uncertainties in these intensities were estimated from duplicate measurement points (T_1, T_2) and the root-mean-squared noise of the matrix base-plane ($\{^1H\}-^{15}N$ NOE). Longitudinal and transverse NMR relaxation decay curves were fit to a single-exponential model, and residual χ^2 values were generally less than the number of fitted points, suggesting robust data analysis. $\{^1H\}-^{15}N$ NOEs were calculated from peak intensity ratios.

^{15}N NMR relaxation parameters were analyzed using the simple Model Free treatment of Lipari and Szabo (*34*, *35*). Here, the spectral density function for isotropic global tumbling is written as:

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_m}{1+\omega^2{\tau_m}^2} + \frac{(1-S^2)\tau}{1+\omega^2\tau^2}\right]$$

where $\tau^{-1} = {\tau_m}^{-1} + {\tau_e}^{-1}$, τ_e is the effective correlation time for internal motion and S^2 is the squared generalized order parameter. Values for ^{15}N bond length and chemical shift anisotropy were taken to be 1.02 Å and −170 ppm, respectively.

The axial symmetric diffusion tensor was determined using a data set trimmed as described below. Analysis resulted in the following parameters: $D_{iso} = 1.48 \times 10^7\ s^{-1}$, $D_{||} = 1.58 \times 10^7$, $D_{\perp} = 1.44 \times 10^7$, $\theta = 1.71$, and $\varphi = 4.43$. The $D_{||}/D_{\perp}$ ratio is then 1.1 for the free protein. This small amount of anisotropy does not change model selection or the overall pattern of Model Free values; therefore, isotropic tumbling was used for final analysis.

The fit of the global correlation time (τ_m) was accomplished using a data set trimmed to exclude data from residues which might exhibit conformational exchange (R_{ex}),

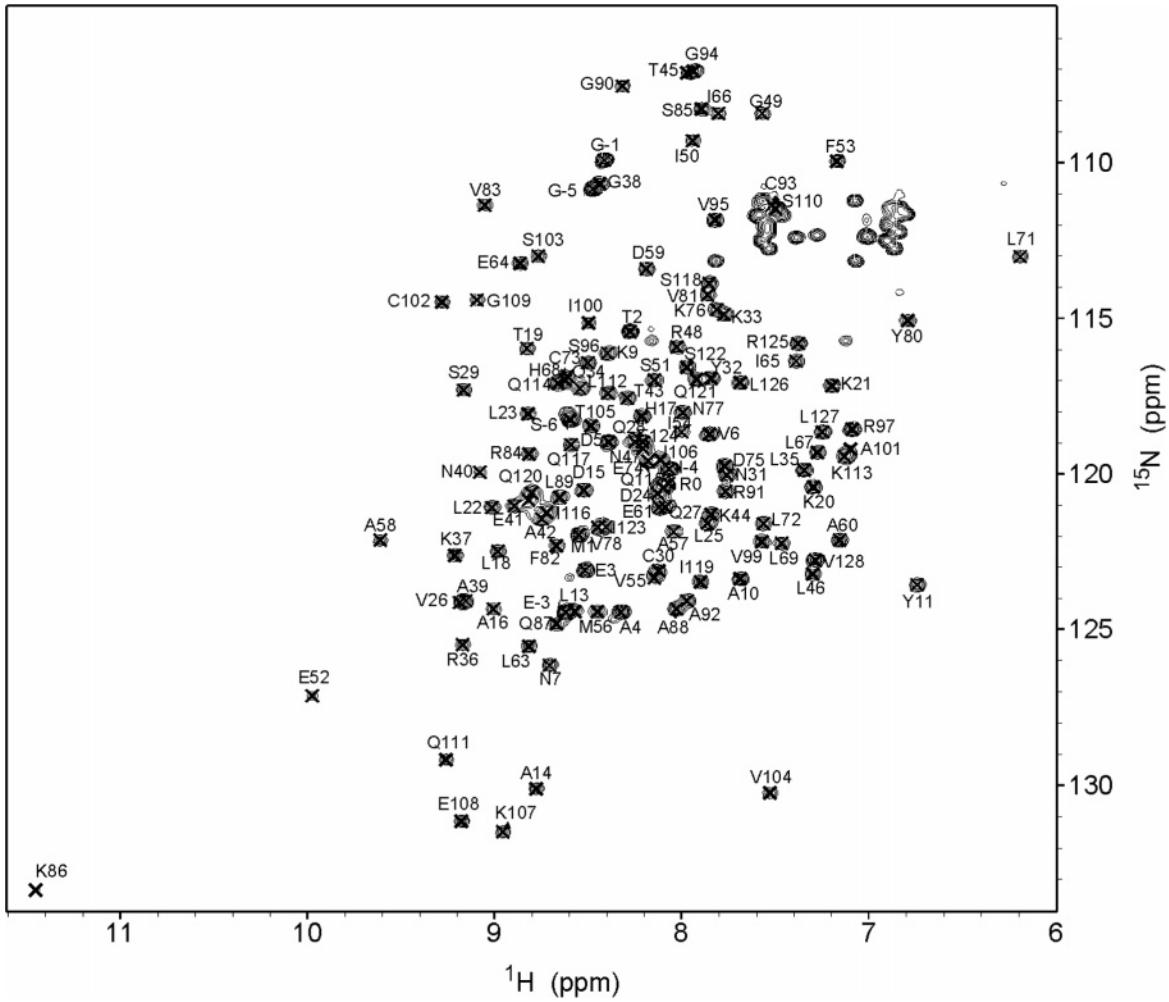

FIGURE 1: ^{15}N-HSQC of 15.5K at 800 MHz in 100 mM sodium chloride, 100 mM sodium phosphate, pH 6.0, at 20 °C shown labeled with resonance assignments.

e.g. those with low NOE values. Data for each residue were then fit using model selection for the five models defined by Palmer and co-workers (*36*). Based on this analysis, each residue was assigned to one of two models such that each residue was fit for S^2 and τ_e, and a few residues included an additional fuj factor such that $R_{ex} = \text{fuj}/(B^2)$ where B is the Larmor frequency in MHz (*37*). The resulting R_{ex} term was only reported where the data could not be well fit by fewer parameters and a significantly positive value resulted. Measured relaxation rates and order parameters have been deposited under accession number 15445 at the Biological Magnetic Resonance Bank.

H/D Exchange. H/D exchange was initiated by buffer exchange into D_2O buffer containing 100 mM sodium chloride, 100 mM sodium phosphate, pD 6.52. The sample was applied to a desalting column equilibrated with D_2O buffer to exchange buffers. The dead time until the midpoint of the first experiment was 11.43 min. A series of 2D ^{1}H−^{15}N HSQC experiments were recorded with measurements taken at intervals over 48 h following initial buffer exchange. Peak intensities were fit to a single-exponential decay function, and exchange rates were converted to protection factors as per Bai et al. (*38*, *39*).

RESULTS

Solution Structure. Solution NMR experiments demonstrate that the 15.5K protein is uniquely folded and stable in solution, as shown in Figure 1. The protein was observed to aggregate slowly in low salt buffers but was found to be stable in 100 mM sodium chloride, 100 mM sodium phosphate, pH 6.0, at 20 °C for several weeks. Backbone amide assignments are also shown in Figure 1, complete resonance assignments have been deposited in the Biological Magnetic Resonance Bank under accession number 7249, and associated information is available (*30*).

The solution structure of 15.5K was calculated using CYANA (*28*, *29*) as described in experimental procedures, and relevant statistics are presented in Table 1. The backbone rmsd for the structured portion of the protein (residues 9−127) of the 20 members of the structure ensemble is 1.18 Å, and the structure has been deposited in the Protein Data Bank archive (*40*) under identification number 2JNB. The rmsd for all heavy atoms was found to be 1.7 Å. The 15.5K structures determined in two previous studies of the protein in binary (15.5K−U4SL) and ternary (15.5K−U4Sl−61K) complexes were refined to 2.9 Å and 2.6 Å, respectively; therefore, all solved structures of 15.5K are of modest resolution. As can be seen in Figure 2, the protein is an α−β−α sandwich as are all other examples of this class. Overall, the protein appears to be well-defined and uniquely structured. The N-terminal affinity tag of the construct was found to be unstructured and thus is not included in the structural data submitted. Additionally, the N-terminus of

Table 1: Statistics for the NMR Solution Structure Ensemble of the Human 15.5K Protein, Submitted to the PDB as 2JNB

NOE distance constraints	
total	1090
short range	713
medium range ($1 < \|i - j\| < 5$)	148
long range ($\|i - j\| \geq 5$)	229
torsion angle restraints:	
total	200
phi	100
psi	100
CYANA target function	0.91
backbone rmsd to mean	1.18 Å
heavy atom rmsd to mean	1.68 Å
Ramachandran analysis:	
most favored	82.3%
additionally allowed	16.6%
generously allowed	1.0%
disallowed	0.0%

the protein up to and including N7 is flexible and unstructured.

The NMR structural model with the least violations was used for comparison with the crystallographic X-ray structures of the complex (PDB IDs 1E7K and 2OZB). The overlay of the three structures is shown in Figure 3, and the alignment of backbone atoms has an rmsd of 1.1 Å for either bound form compared to the free structure, which is within the rmsd between individual members of the structure ensemble. The two structures of bound forms of 15.5K are nearly identical. Differences in the relative placement of helices between the free and bound forms of 15.5K include the rotation of α4 away from the protein and a slight curve in α5 resulting in new position for the C-terminus. The most notable difference between the crystal structure of the complex and the solution structure presented here occurs for residues 63−78 (α3). Analysis of chemical shifts in the α3 region supports the secondary structure pattern seen previously of a short helix (63−65) followed by a much longer helix (67−75) (*23*); however, the NMR structure calculations resulted in a single helical element which is capped by P70 (see Figures 2 and 3). Medium-range NOEs indicative of helices were not found for residues 70−76. In fact, relatively few restraints were found for this region, which is consistent with conformational fluctuations.

^{15}N NMR Relaxation Analysis. Knowledge of the dynamic characteristics of a protein is critical to understanding its structure−function relationship. Dynamics constitute a fundamental mechanism by which proteins can explore multiple conformations in the search for a structure that either recognizes a target structural motif or presents a recognition motif to a binding partner. Additionally, biological function often depends on system transitions that proceed from states of lower energy to states of higher energy, which may be accessed via local fluctuations. Longitudinal (T_1) and transverse (T_2) ^{15}N NMR relaxation together with the $\{^1H\}-^{15}N$ NOE yield information about both global tumbling and the motion of backbone H−N bond vectors. Derived bond vector motion on the relatively fast ps time scale is accompanied with information about the motional amplitude. ^{15}N relaxation can also report on motion on the μs to ms time scale, and although motional amplitudes cannot be derived in this regime, under favorable circumstances the relative population of states in exchange may be obtained (*41*). In addition to the sub-ns and μs to ms time scales, motion occurring on the time scale from $\sim 10^2$ ps to a few ns can be detected as a quenching of the anticipated NOE, e.g., as evaluated in regular secondary structure elements wherein large NOE values are expected.

The relaxation analysis for the free 15.5K protein used 99 residues for which the resonances could be clearly resolved. The initial data set containing T_1, T_2, and NOE measurements for the two fields was then trimmed as described in Experimental Procedures to remove residues exhibiting conformational exchange. The global fit of the trimmed data set for the combined fields resulted in a global correlation time of $\tau_m = 11.75$ ns, which is reasonable for these conditions and the estimated volume of the protein. Analysis of the axial symmetry of the protein revealed an axial ratio ($D_\parallel/D_\perp$) of 1.1 (*42*). Consideration of the fitted motional parameters confirms that this minor anisotropy does not change either selection of the optimal model nor does it significantly influence the determined S^2, τ_e, or R_{ex} values (*43*). Therefore, isotropic tumbling was assumed for data analysis and the resulting Model Free parameters for free 15.5K in solution are displayed in Figure 4. Data were fit to one of two models (with or without R_{ex}) based on a combination of model-fitting analyses as implemented in relxn2.2 (*37*). Residue E52 could not be fit to any model with reasonable values. Apart from the flexible N-terminal region, squared generalized order parameters (S^2) are uniformly high, $\langle S^2 \rangle \sim 0.9$, indicating that the protein backbone is relatively rigid on the subnanosecond time scale. Several residues exhibit relatively low order parameters, which suggest regions of local flexibility in the fast-motion time scale (ps to ns). These are highlighted in yellow in Figure 5 and correspond to both termini and two residues found to interact with the U4SL, E41, and R97 (*23*). The correlation times for local backbone motion display a similar pattern such that the bulk of the residues exhibit relatively rapid local correlation times ($\langle \tau_e \rangle \sim 62$ ps).

The R_{ex} component of relaxation depends in a complex fashion on the relative populations of subconformers (p_A and p_B), the related chemical shift differences of nuclei in the various conformational states ($\Delta\omega_N$), and the rate of interconversion between such states (k_{ex}) (*44*). The R_{ex} values determined in the current study are relatively small, and attempts to directly measure the R_{ex} contribution using an R_2-CPMG experiment (*45*, *46*) were unsuccessful (see Figure S1 of the Supporting Information). Therefore, inasmuch as the focus of the current study is aimed at elucidating the salient features of the dynamics of the free protein, for the purposes of the current analysis, the presence or absence of relatively slow (μs to ms) local motions is the key observation.

The relaxation model evaluation process described in experimental procedures identified 17 residues with small but statistically significant R_{ex} contributions to relaxation. Residues were only fit to a model including R_{ex} if the data were not well fit to a simpler model, and if the result of including such a term resulted in a significantly positive value. These are identified in red in context of the sequence and secondary structure in Figure 5. As can be seen in this figure, these regions are localized to α-helices 1, 4, and 5.

Hydrogen Exchange of Amides (HX). The exchange rates of amide hydrogens have long been used as a probe of

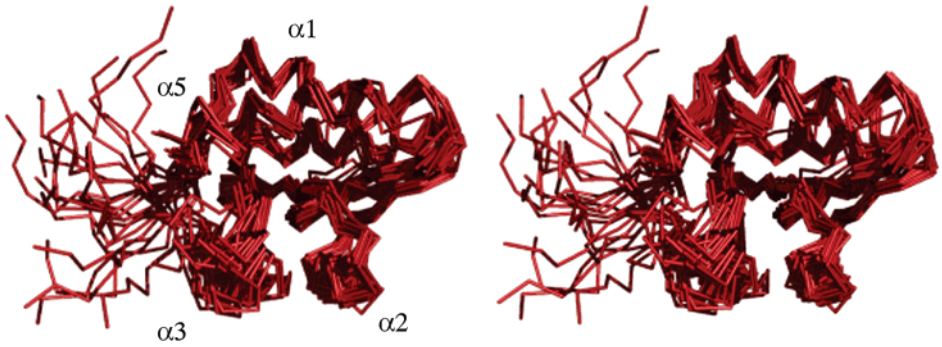

FIGURE 2: Solution NMR structure ensemble of 15.5K with an rmsd of 1.68 Å. The stereoview image pair shows the ensemble of 20 structures, well-defined aside from the N-terminus and α3 regions.

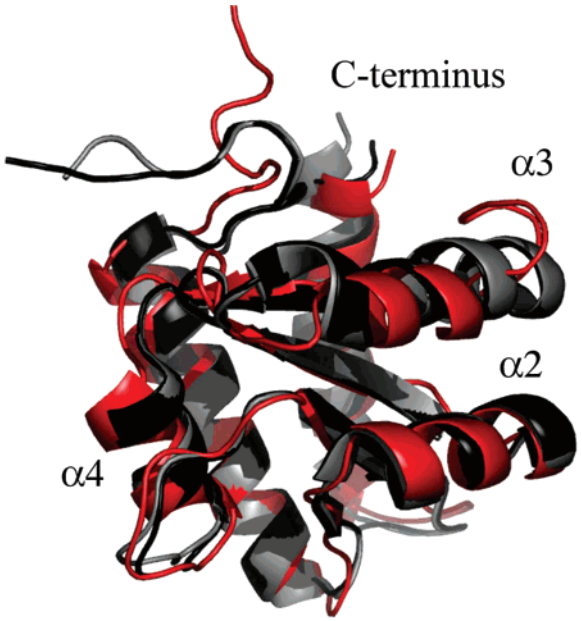

FIGURE 3: Agreement between structures and states of 15.5K shown by overlay of a representative structure from the NMR structure ensemble (red), isolated chain A from PDB entry 1E7K (gray), and isolated chain A from PDB entry 2OZB, (black). The latter two are almost identical structures. The structures are oriented such that the U4 RNA binding interface is at the bottom of the protein.

solvent accessibility and motion in proteins (*47*, *48*). Experimentally measured exchange rates are compared to those determined for random coil peptides and converted into slowing or protection factors (*39*). The level of protection is due to contributions from both static and transient exposure to the solvent. The data is interpreted in terms of the protein structure such that regular secondary structures would be expected to be more protected, whereas the regions lying outside of regular structural elements will experience less protection. In addition, these exchange rates can result from a wide range of timescales and generate a view of internal protein motion which complements that revealed by relaxation analysis.

HX rates were measured for the native state of free 15.5K in solution at 20 °C. The most rapid rate that could be measured using the experimental scheme was ~0.09 min^{-1}. The fitted rates are presented along with the log of the protection factors ($\log(k_c/k_{ex})$) as Supporting Information. These protection factors are plotted relative to secondary structure elements in Figure 6. Analysis of this data reveals regions of the protein which appear less protected than would

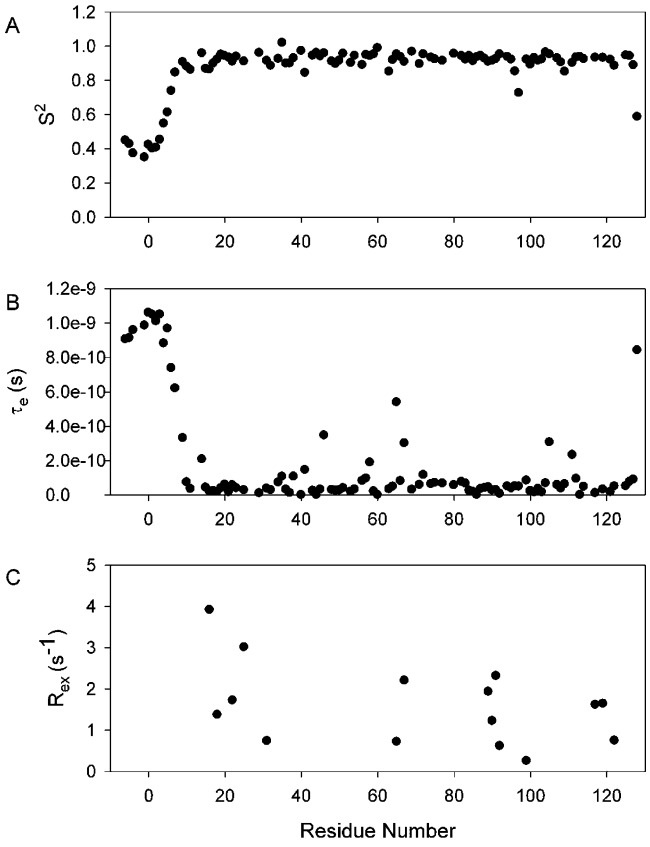

FIGURE 4: Model Free analysis of backbone NMR relaxation data for 15.5K free in solution at 20 °C. Data recorded at 500 and 600 MHz (^{1}H) were used to fit (a) S^2, (b) τ_e, and (c) R_{ex} for $\tau_m = 11.75$ ns as described in Experimental Procedures.

be otherwise expected based on consideration of the surrounding residues and secondary structure elements. Some of these residues correspond to the K-turn RNA binding interface, including V95, R97, and I100. Other significant regions of low protection include residues 86−88 of helix 4 and the α3 region, which will be discussed in greater detail below.

DISCUSSION

Structural changes that occur as a result of the interactions between proteins and RNA oligonucleotides are of intrinsic importance due to their vital functional roles. Specific examples of K-turn binding proteins exhibit distinct patterns of structure and dynamics that may be associated with differences in their precise functional roles. L7Ae is a member of the K-turn RNA binding protein family and the

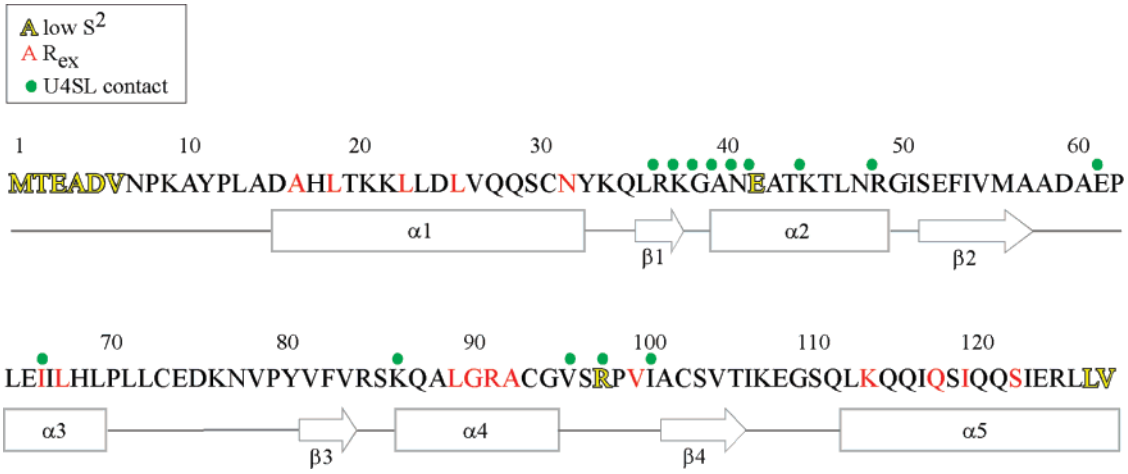

FIGURE 5: Sites of conformational exchange or flexibility for free 15.5K in solution displayed in context of the secondary structure and protein−RNA contacts. Residues demonstrating R_{ex} for the backbone ^{15}N nucleus are highlighted in red, residues which are flexible on the basis of a low S^2 in yellow, and U4SL contacts are highlighted in green as described in Vidovic et al., 2000.

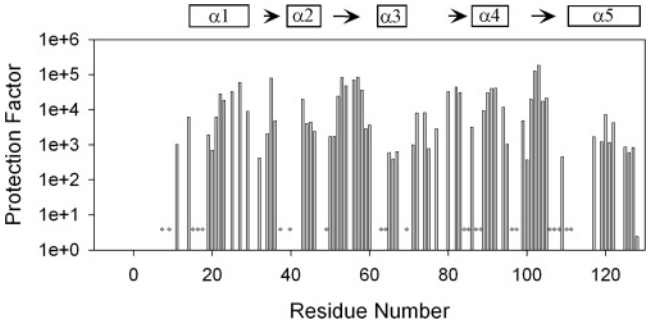

FIGURE 6: H/D exchange protection factors for 15.5K in solution at 20 °C are displayed alongside secondary structure elements. Residues for which exchange was too rapid to be measured are indicated by an asterisk.

functional homologue of 15.5K in Archaea with 33% sequence identity to 15.5K (*49*). This protein binds K-turns found in the ribosome and the sRNA box C/D and box H/ACA motifs, and its structure has been studied in complex with various targets (*13*, *16*, *17*). Recently, a structure of free L7Ae was also determined, in which the backbone of the free protein structure shows no significant deviation from bound L7Ae, which suggests that the recognition and binding of the K-turn does not entail a significant structural conformation change for the protein (*18*). Additional new structural data for Snu13p (70% identical to 15.5K) and *M. jannaschii* L7Ae lend further support to this conclusion (*15*, *19*). These structures are in contrast to the structural differences previously seen for L30, which has likewise been characterized in both its free and bound forms (*20*−*22*). The ribosomal L30 protein is not a functional homologue of 15.5K, but is clearly a member of the same structural family of K-turn binding proteins and retains 18% sequence identity (*20*). In the eukaryotic L30, a segment of the protein corresponding to residues 86−100 in 15.5K undergoes secondary structure transitions upon complex formation (*21*). Interestingly, the free L30 was only stable at higher salt concentrations and low temperatures, as we have seen for 15.5K, while the L30/pre-mRNA complex was stable at low salt and room temperature, indicating the apparent stabilizing effect that RNA binding has on the protein (*20*). In addition, a number of domains of the free L30 protein appear to undergo conformational exchange but adopt a unique conformation in the complex, which manifests itself as the sharpening of resonances in the 2D ^{15}N-HSQC spectrum (*20*).

The bulk of backbone amide residues in free 15.5K exhibit relatively high order parameters and rapid local correlation times, and relaxation data for the bulk of the molecule are well-fit using the simple Model Free analysis. The implication of this result, with respect to an induced fit hypothesis for the interaction of 15.5K and its RNA target, is thus that 15.5K presents a relatively rigid initial contact platform on the timescales sampled by NMR relaxation analysis. Although the results of ^{15}N NMR relaxation analysis confirm that the protein is relatively rigid, departures from this general trend are found at the N-terminus and several isolated sites. The N-terminus through V6 is flexible or unstructured and does not exhibit regular secondary structure until A16. On the basis of chemical shift analysis, residues 12−15 appear to have some β-strand-like character, but no NOEs were found to support including this region as part of the core β-sheet.

Residues 63−78 of 15.5K are somewhat underconstrained in the structural analysis; however, some aspects of this region of the protein may be illuminated by discussion of the observations available. Chemical shift analysis (not shown) predicts a helical structure, but NOEs to unambiguously support this classification for the entire region were not found. Structure calculations suggest that a single helix spans the first half of the region, is capped by P70, and continues as random coil. ^{15}N relaxation analysis shows no unusual features other than potential RNA-binding residues discussed below. On the other hand, the H/D exchange data available reveal a lack of protection from exchange throughout the segment. The lack of robust NOE constraints and the relatively weak protection against hydrogen exchange, taken together, indicate that portions of the α3 helix may undergo conformational fluctuations. Complementary data from a recent mutational survey of surface residues in 15.5K reveals that residues 74−77 are critically important for binding the 61K protein (*50*). The 15.5K−61K interaction is believed to be essential for U4/U6 assembly and remodeling, and the involvement of the α3 region of 15.5K in this interaction has been confirmed by a ternary structure of 15.5K−U4−61K (*51*). This information provides a compelling context for interpreting the NMR relaxation and protec-

tion data and suggests that local motion present in α3 may be an important attribute of the protein interaction interface of 15.5K.

The structure and dynamics of the α4 are of special interest due to its homology to a mobile loop in the L30 protein, which also binds K-turn RNAs and has been shown to undergo structural rearrangement upon binding. The major rearrangement observed in L30 is the conversion of random coil to helical structure in the region spanning residues 74−81 (*21*). In 15.5K the corresponding helix consists of residues 86−93, which then leads into a loop pointing out toward the RNA interface. Although the segment has a unique helical structure in the free protein, the dynamical properties are interesting. The N-terminus of the helix is weakly protected from hydrogen exchange including two residues with rates too fast to measure. Residues 89−92 exhibit conformational exchange in the free solution state, yet the amide protons are highly protected from solvent exchange. Therefore, although the structure of the helix is well-defined, there are unique local motional factors which have a potential functional role for the protein.

The RNA-binding surface of 15.5K is clearly of central importance to the current study. Analysis of the RNA-binding surface suggests that regions of the protein that contact the RNA possess unique structural and dynamical features, for which several key positions are discussed below. I65, which is in the middle of α3, exhibits R_{ex} and minimal protection from HX, and E41, at the beginning of α2, is more flexible (low S^2). Other notable positions are localized to the loop spanning residues 95−101, wherein residues V95, R97, and I100 show limited protection from hydrogen exchange events, and marked flexibility is noted for R97. This loop is of particular interest because of its homology to the mobile loop in L30 (*21*) and its position on the RNA binding interface of 15.5K. Therefore, conformational exchange and other fast motions are evident in strategic locations that are likely to influence adaptive interactions between the 15.5K and its RNA targets.

In summary, the mechanism of protein−RNA recognition for the K-turn systems such as 15.5K has important implications for the general role of induced fit in recognition and binding for protein−RNA complexes. The eukaryotic L30 protein has been shown to undergo structural rearrangement whereas the 15.5K orthologues Snu13p (yeast) and L7Ae (Archaea) apparently do not, and recent work has demonstrated significant conformational flexibility in K-turn RNAs (*14*, *52*−*54*). Results of the current study support the theory that induced fit drives the 15.5K−U4SL interaction through localized motions in the protein combined with presumably larger conformational changes in the structure of the RNA. Backbone dynamics and H/D exchange data add further support to minor structural changes in some regions of the protein upon binding RNA. The work described here is the first study of the dynamics of a eukaryotic K-turn RNA binding protein free in solution and confirms the presence of several regions of conformational exchange or flexibility in the free protein. The high level of sequence and structural similarity among 15.5K orthologues also establishes the importance and applicability of the results of the dynamical studies discussed here.

ACKNOWLEDGMENT

The authors would like to thank Dr. Jack Skalicky of the Department of Biochemistry, University of Utah for assistance in structure analysis, and Dr. Andrew Lee of the Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, for the updated relaxation analysis software. Portions of this work were conducted at the Health Sciences NMR Facility at the University of Utah. NOESY data were collected at the W. M. Keck High Field NMR Facility at the University of Colorado with assistance from Dr. Andrew Fowler. Backbone resonance assignment data were collected with the assistance of Dr. Nancy Isern at the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory.

SUPPORTING INFORMATION AVAILABLE

A table containing the observed hydrogen exchange rates and subsequent protection factors for free 15.5K, and example plots for R_{ex} measurements for 15.5K are also presented. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Moore, M. J., Query, C. C., and Sharp, P. A. (1993) in *The RNA world* (Gesteland, R. F., and Atkins, J. F., Eds.) pp 303−357, Cold Spring Harbor Laoratory Press, Cold Spring Harbor, NY.
2. Madhani, H. D., and Guthrie, C. (1994) Dynamic RNA-RNA interactions in the spliceosome, *Annu. Rev. Genet. 28*, 1−26.
3. Staley, J. P., and Guthrie, C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things, *Cell 92*, 315−326.
4. Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R., and Luhrmann, R. (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5′ stem-loop of U4 snRNA, *EMBO J. 18*, 6119−6133.
5. Marmier-Gourrier, N., Clery, A., Senty-Segault, V., Charpentier, B., Schlotter, F., Leclerc, F., Fournier, R., and Branlant, C. (2003) A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on the U3 small nucleolar RNA, *RNA 9*, 821−838.
6. Dobbyn, H. C., and O'Keefe, R. T. (2004) Analysis of Snu13p mutations reveals differential interactions with the U4 snRNA and U3 snoRNA, *RNA 10*, 308−320.
7. Watkins, N. J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., and Branlant, C. (2000) A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP, *Cell 103*, 457−466.
8. Gottschalk, A., Neubauer, G., Banroques, J., Mann, M., Luhrmann, R., and Fabrizio, P. (1999) Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP, *EMBO J. 18*, 4535−4548.
9. Nottrott, S., Urlaub, H., and Luhrmann, R. (2002) Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins, *EMBO J. 21*, 5527−5538.
10. Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs, *EMBO J. 20*, 3617−3622.
11. Maxwell, E. S., and Fournier, M. J. (1995) The small nucleolar RNAs, *Annu. Rev. Biochem. 35*, 897−934.
12. Watkins, N. J., Dickmanns, A., and Luhrmann, R. (2002) Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5k protein, for the hierarchical assembly of the box C/D snoRNP, *Mol. Cell. Biol. 22*, 8342−8352.
13. Klein, D. J., Schmeing, T. M., and Moore, P. B. (2001) The kink-turn: a new RNA secondary structure motif, *EMBO J. 20*, 4214−4221.

14. Goody, T. A., Melcher, S. E., Norman, D. G., and Lilley, D. M. J. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependant, *RNA 10*, 254−264.
15. Oruganti, S. v., Zhang, Y., and Li, H. (2005) Structural comparison of yeast snoRNP and spliceosomal protein Snu13p with its homologs, *Biochem. Biophys. Res. Commun. 333*, 550−554.
16. Moore, T., Zhang, Y., Fenley, M. O., and Li, H. (2004) Molecular basis of box C/D RNA-protein interactions: cocrystal structure of archaeal L7Ae and a box C/D RNA, *Structure 12*, 807−818.
17. Hamma, T., and Ferre-D'Amare, A. R. (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution, *Structure 12*, 893−903.
18. Charron, C., Manival, S., Clery, A., Senty-Segault, V., Charpentier, B., Marmier-Gourrier, N., Branlant, C., and Aubry, A. (2004) The archeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity, *J. Mol. Biol. 342*, 757−773.
19. Suryadi, J., Tran, E. J., Maxwell, E. S., and Brown, B. A. I. (2005) The crystal structure of the Methanocaldococcus jannaschii multifunctional L7Ae RNA-binding protein reveals an induced-fit interaction with the box C/D/RNAs, *Biochemistry 44*, 9657−9672.
20. Mao, H., and Williamson, J. R. (1999) Local folding coupled to RNA binding in the yeast ribosomal protein L30, *J. Mol. Biol. 292*, 345−359.
21. Chao, J. A., Prasad, G. S., White, S. A., Stout, C. D., and Williamson, J. R. (2003) Inherent protein structural flexibility at the RNA-binding interface of L30e, *J. Mol. Biol. 326*, 999−1004.
22. Chao, J. A., and Williamson, J. R. (2004) Joint x-ray and NMR refinement of the yeast L30e-mRNA complex, *Structure 12*, 1165−1176.
23. Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R., and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment, *Mol. Cell 6*, 1331−1342.
24. Williamson, J. R. (2000) Induced fit in RNA-protein recognition, *Nat. Struct. Biol. 7*, 834−837.
25. Chen, Y., and Varani, G. (2005) Protein families and RNA recognition, *FEBS J. 272*, 2088−2097.
26. Goddard, T. D., and Kneller, D. G., University of California, San Fransisco.
27. 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 a phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation, *Biochemistry 33*, 5984−6003.
28. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA., *J. Mol. Biol. 273*, 283−298.
29. Hermann, T., Guntert, P., and Wuthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *J. Mol. Biol. 319*, 209−227.
30. Soss, S. E., and Flynn, P. F. (2007) NMR assignment of the human spliceosomal 15.5K, *J. Biomol. NMR 38*, 175.
31. Cornilescu, G., Delaglio, F., and Bax, A. (1999) TALOS: protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.54*, 905−921.
33. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR., *J. Biomol. NMR 8*, 477−496.
34. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546−4559.
35. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in Macromolecules 2. Analysis of experimental results, *J. Am. Chem. Soc. 104*, 4559−4570.
36. Mandel, A. M., Akke, M., and Palmer, A. G., III. (1995) Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme, *J. Mol. Biol. 246*, 144−163.
37. Clarkson, M. W., Gilmore, S. A., Edgell, M. H., and Lee, A. L. (2006) Dynamic coupling and allosteric behavior in a nonallosteric protein, *Biochemistry 45*, 7693−7699.
38. Connelly, G. P., Bai, Y., Jeng, M.-F., and Englander, S. W. (1993) Isotope effects in peptide group hydrogen exchange, *Proteins 17*, 87−92.
39. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange, *Proteins 17*, 75−86.
40. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
41. Palmer, A. G. (2004) NMR characterization of the dynamics of biomacromolecules, *Chem. Rev. 104*, 3623−3640.
42. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation, *J. Am. Chem. Soc. 117*, 12562−12566.
43. Schurr, J. M., Babcock, H. P., and Fujimoto, B. S. (1994) A test of the Model-Free formulas. Effects of anisotropic rotational diffusion and dimerization, *J. Magn. Reson. B105*, 211−214.
44. Palmer, A. G., III, Williams, J., and McDermott, A. (1996) Nuclear magnetic resonance studies of biopolymer dynamics, *J. Phys. Chem. 100*, 13293−13310.
45. Massi, F., Grey, M. J., and Palmer, III, A. G. (2005) Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1p relaxation experiments, *Protein Sci. 14*, 735−742.
46. Loria, J. P., Rance, M., and Palmer, III, A. G. (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy, *J. Am. Chem. Soc. 121*, 2331−2332.
47. Linderstrom-Lang, K., and Schellman, J. A. (1959) in *The Enzymes* (Lardy, H., and Myrback, K., Eds.) pp 443−510, Academic Press, New York.
48. Englander, S. W., and Kallenbach, N. R. (1984) Hydrogen exchange and structural dynamics of proteins and nucleic-acids, *Q. Rev. Biophys. 16*, 521−655.
49. Kuhn, J. F., Tran, E. J., and Maxwell, E. S. (2002) Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein, *Nucleic Acids Res. 30*, 931−941.
50. Schultz, A., Nottrott, S., Watkins, N. J., and Luhrmann, R. (2006) Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs, *Mol. Cell. Biol. 26*, 5146−5154.
51. Liu, S., LI, P., Dybkov, O., Nottrott, S., Hartmuth, K., Luhrmann, R., Carlomagno, T., and Wahl, M. C. (2007) Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP, *Science 316*, 115−119.
52. Matsumura, S., Ikawa, Y., and Inoue, T. (2003) Biochemical characterization of the kink-turn RNA motif, *Nucleic Acids Res. 31*, 5544−5551.
53. Cojocaru, V., Nottrott, S., Klement, R., and Jovin, T. M. (2005) The snRNP 15.5K protein folds its cognate K-turn RNA: A combined theoretical and biochemical study, *RNA 11*, 197−209.
54. Razga, F., Koca, J., Sponer, J., and Leontis, N. B. (2005) Hinge-like motions in RNA Kink-turns: the role of the second A-minor motif and nominally unpaired bases, *Biophys. J. 88*, 3466−3485.

BI701254Q