Structure of the Nuclear Factor ALY: Insights into Post-Transcriptional Regulatory and mRNA Nuclear Export Processes†

Gabriela C. Pérez-Alvarado,‡ Maria Martínez-Yamout,‡ Melissa M. Allen,‡ Rudolf Grosschedl,§ H. Jane Dyson,‡ and Peter E. Wright*,‡

Department of Molecular Biology and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, and Gene Center and Institute of Biochemistry, University of Munich, 81377 Munich, Germany

Received January 13, 2003; Revised Manuscript Received April 18, 2003

ABSTRACT: ALY is a ubiquitously expressed nuclear protein which interacts with proteins such as TAP that are involved in export of mRNA from the nucleus to the cytoplasm, as well as with proteins that bind the T cell receptor α gene enhancer. ALY has also been shown to bind mRNA and to co-localize in the nucleus with components of a multiprotein postsplicing complex that is deposited 20−24 nucleotides upstream of exon−exon junctions. ALY has a conserved RNA binding domain (RBD) flanked by Gly-Arg rich N-terminal and C-terminal sequences. We determined the solution structure of the RBD homology region in ALY by nuclear magnetic resonance methods. The RBD motif in ALY has a characteristic $\beta_1\alpha_1\beta_2-\beta_3\alpha_2\beta_4$ fold, consisting of a β sheet composed of four antiparallel β strands and two α helices that pack on one face of the β sheet. As in other RBD structures, the β sheet has an exposed face with hydrophobic and charged residues that could modulate interactions with other molecules. The loop that connects β strands 2 and 3 is in intermediate motion in the NMR time scale, which is also characteristic of other RBDs. This loop presents side chains close to the exposed surface of the β sheet and is a primary candidate site for intermolecular interactions. The structure of the conserved RBD of ALY provides insight into the nature of interactions involving this multifunctional protein.

ALY (ally of AML-1 and LEF-1 proteins)[1] (*1*), also known as REF1-I (RNA and export factor binding-I) (*2*) and BEF (bZIP enhancing factor) (*3*), is a nuclear protein which was originally proposed to play an essential role in the activation of the T-cell receptor α gene (TCRα) enhancer by interacting with the transcription factors AML-1 and LEF-1 (*1*). Other studies have shown that ALY exhibits chaperone function. Through direct interaction with basic region-leucine zipper (bZIP) transcription factors, ALY enhances binding of bZIP factors to DNA (*3*). Since the TCRα enhancer contains a binding site for the ATF family of bZIP proteins, ALY may be an important factor in facilitating the formation of a higher order nucleoprotein complex on the TCRα enhancer.

ALY is a member of the REF family of proteins and interacts with the mRNA export factor TAP, facilitating its interaction with cellular mRNAs (*2*, *4*). ALY itself exhibits RNA binding activity and may act as an RNA chaperone by modulating the formation of RNA−protein and RNA−RNA complexes (*2*, *4*). Recently, it was reported that ALY is a component of the general mRNA export machinery (*5*). ALY co-localizes with splicing factors in the nucleus, shuttling between the nucleus and the cytoplasm to increase the rate and efficiency of mRNA export (*6*). ALY is found in a multiprotein postsplicing complex, also known as the exon junction complex (EJC) (*7*) with SRm160, DEK, RNPS1, Y14, and Magoh (*8*, *9*). ALY interacts with pre-mRNA prior to spliceosome assembly, and during or after spliceosome formation, Y14, DEK, SRm160, and RNPS1 associate at the −20 to −24 region of the 5′ exon and remain stably bound to mRNA without apparent preference for a particular RNA sequence (*7*, *8*). This postsplicing complex tags mRNAs for nuclear export and surveillance during which mRNAs with truncated open reading frames are detected and subjected to nonsense-mediated mRNA decay. Surveillance of mRNA is promoted by interaction with hUpf3, which joins the complex via interaction with RNPS1 (*10*) and Y14 (*11*). hUpf3 may become associated with mRNA at a late post-splicing stage. Coimmunoprecipitation studies have shown that hUpf3, ALY, and TAP are found within the same complex prior to mRNA export (*11*). Export is promoted by the recruitement of TAP, by either ALY and/or Y14, before mRNA export. After export to the cytoplasm, ALY and TAP are dissociated, and Y14 (*8*), Magoh (*9*), RNPS1, and hUpf3b (*10*) remain bound to the mRNA upstream from the exon-exon junction site.

† This work was supported by Grants GM36643 and CA96865 from the National Institutes of Health and by the Skaggs Institute for Chemical Biology. G.C.P.A. acknowledges support from the Leukemia and Lymphoma Society.

* Corresponding author. Telephone: (1) (858)-784-9721. Facsmile: (1) (858)-784-9822. E-mail: wright@scripps.edu.

‡ The Scripps Research Institute.

§ University of Munich.

[1] Abbreviations: AML1, acute myeloid leukemia 1; LEF-1, lymphocyte enhancer-binding factor; TAP, tip associating protein; NMR, nuclear magnetic resonance; HSQC, heteronuclear single quantum coherence; 2D, two-dimensional; 3D, three-dimensional; NOESY, nuclear Overhauser effect spectroscopy; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate, sodium salt.

10.1021/bi034062o CCC: $25.00
Published on Web 05/30/2003

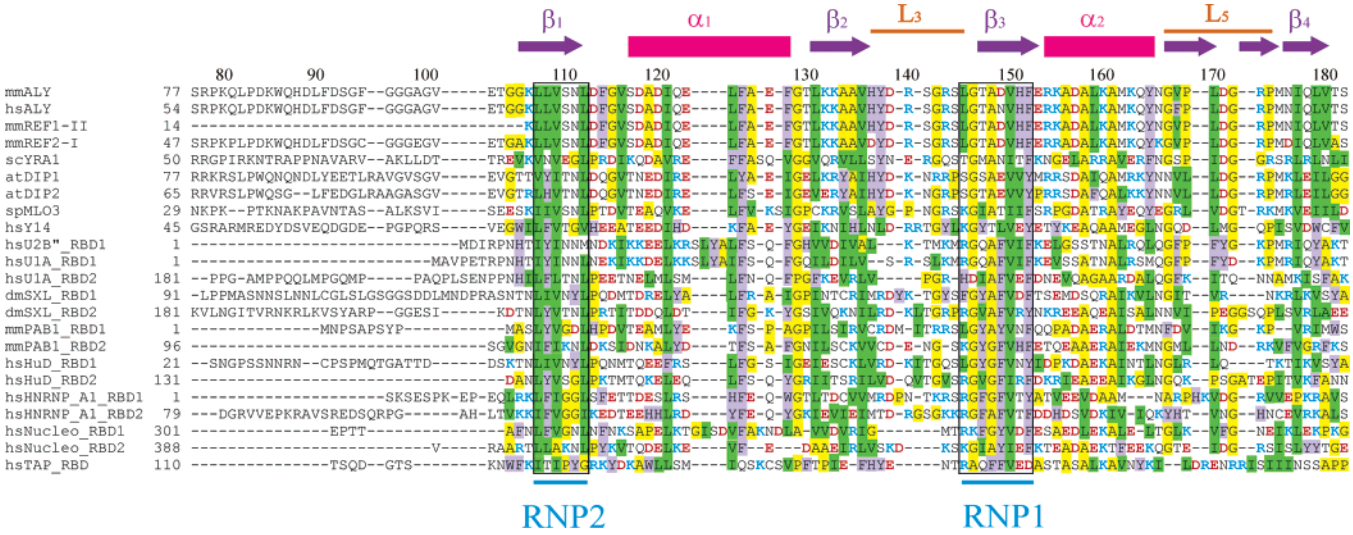

FIGURE 1: Multiple sequence alignment of the RBD domain of mouse ALY/mREF1-I (*1*) and flanking sequences against the corresponding sequences of human (hs-*Homo sapiens*) ALY (*76*), mouse (mm-*Mus musculus*) REF1-II (*2*), mouse REF2-I (*2*), yeast (sc-*Saccharomyces cerevisiae*) YRA1 (*77*), *Arabidopsis thaliana* (at) DIP1 and DIP2 (*71*), *Schizosaccharomyces pombe* (sp) MLO3 (*72*), human Y14 (*78*), human U2B″-RBD1 (*63*), human U1A-RBD1 (*55*−*58*) and RBD2 (*59*), *Drosophila melanogaster* (ds) SXL-RBD1 (*60*, *61*) and RBD2 (*61*, *62*), mouse PABP1-RBD1 and RBD2 (*64*), human HuD-RBD1 and RBD2 (*70*), human hnRNPA1-RBD1 (*67*−*69*) and RBD2 (*68*, *69*), human nucleolin-RBD1 and RBD2 (*65*, *66*), and human TAP-RBD. Secondary structural elements based on the ALY77−182 structure are denoted on the top of the alignment. Strands are in purple and helices in magenta. Black boxes denote the conserved RNP1 and RNP2 submotifs (cyan). The alignment was created from WU-Blast2 alignments and structured-based alignments for sequences with a structure. The residues are highlighted in green (aliphatic), yellow (Ala, Gly, Pro), and purple (aromatic). Charged side chains are denoted in blue (positive) and red (negative).

Y14 binds preferentially to spliced mRNAs and is stably bound to sequences immediately upstream of exon−exon junctions. It appears to provide a position-specific molecular memory that communicates to the cytoplasm the location of exon and intron boundaries. Y14 co-immunoprecipitates with ALY from the nucleus. This complex is resistant to RNase A treatment, suggesting that Y14 and ALY bind to each other through protein−protein interactions instead of through tertiary interactions with mRNA (*12*).

Recent studies have shown that ALY interacts with the conserved DEAD-box helicase UAP56, which functions during spliceosome assembly. It is has been suggested that UAP56 recruits ALY to the spliced mRNA−protein complex coupling splicing and export (*13*). ALY and UAP56 associate with proteins that form part of a transcription/export complex (TREX), which is recruited to activate genes during transcription and travels the entire length of the gene with RNA polymerase II during transcriptional elongation. The TREX complex appears to have a conserved role in coupling transcription to mRNA export (*14*).

Although ALY is clearly important for mRNA splicing and export and also for enhancing transcription activation by several nuclear factors, the details of its function are not well understood. The available evidence suggests that ALY forms an interface for protein−protein interactions and also binds mRNA. As a step toward understanding the function of ALY at the molecular level, we have determined the three-dimensional structure of residues Ser77 to Ser182 of mouse ALY (ALY77−182) encompassing an RBD (RNA binding domain), also known as RNA recognition motif (RRM), or ribonucleoprotein (RNP) domain homology region (Figure 1). The three-dimensional structure provides insights into the potential sites of interaction of ALY and allows us to make some predictions regarding the nature of its specific interactions with either proteins or ribonucleic acids.

MATERIALS AND METHODS

Preparation of Protein Samples. The RBD homology region (residues 77−182) of mouse ALY was cloned into the T7 expression vector pET21a(+) (Novagen, Wisconsin). The protein was expressed in *Escherichia coli* BL21(DE3) cells harboring the vector, as well as the pUBS520 vector containing the dnaY gene ($tRNA^{Arg}_{AGA/AGG}$), at 37 °C in minimal media. Cells were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) between 0.7 and 0.8 OD_{600}, incubated at 25 °C and harvested after 11−12 h when they reached $OD_{600} = 2$. The cell pellet was resuspended in 50 mM Tris-HCl, pH 7, 0.1 M NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 2 mM EDTA (ethylenediamine tetraacetic acid disodium salt), 1% NP-40 (nonaethylene glycol octyphenyl ether), and 100 mM DTT (dithiothreitol). The cells were lysed after sonication at 4 °C and centrifuged, and the supernatant containing ALY77−182 was passed through a 5 mL Hi-Trap Q column (Pharmacia) equilibrated in 10 mM MES buffer, pH 6. The flow through was purified by FPLC using 5 mL Hi-Trap SP columns (Pharmacia), and the target protein was eluted with a NaCl gradient. The eluted fractions were concentrated using a Centriprep-3 (Amicon) and loaded onto a Superdex-75 column (Pharmacia). The purity of the isolated protein was determined by Coomassie gradient SDS−PAGE (Novex, Invitrogen), analytical reversed-phase HPLC, and electrospray mass spectroscopy. Pure fractions were concentrated to 0.5 mM and exchanged into the NMR buffer (20 mM Tris-HCl-d8, pH 6.5, 100 mM NaCl, 0.2 mM EDTA, 0.01%NaN_3 and 10% D_2O) in a Centriprep-3 concentrator (Amicon, Inc.). ALY was uniformly labeled with ^{15}N and/or ^{13}C by growing cells in minimal medium with $(^{15}NH_4)_2SO_4$ or $^{15}NH_4Cl$ and/or $^{13}C_6$-glucose as sole sources of nitrogen and carbon.

NMR Spectroscopy. All NMR spectra were recorded at 23 and 16 °C on Bruker AMX-I and -II 500 MHz, DRX 600 and 800 MHz spectrometers. Data were processed using NMRPipe and NMRDraw (*15*) and analyzed using NMRVIEW (*16*). Backbone resonances of ALY were assigned sequence-specifically from 3D HNCA (*17*), HN(CA)CB (*18*), CBCA(CO)NH (*19*), and HN(CO) (*20*) data. Side chain assignments were determined from 3D C(CO)NH-TOCSY (*21*), HC(CO)NH-TOCSY (*21*), ^{15}N-edited TOCSY (*22*), HCCH-COSY (*23*), CCH-COSY (*24*), HCCH-TOCSY (*23*), 2D-^{13}C-HMQC, DQF-COSY (*25*), and ^{1}H-^{1}H-TOCSY (*26*) spectra. The results from the ^{15}N-NOESY-HSQC (*22*), ^{13}C-NOESY-HSQC (*27*), ^{15}N/^{13}C-NOESY-HSQC (*28*), and ^{13}C-HMQC-NOESY (*29*) spectra allowed determination of the topology of ALY77−182. Secondary structural elements and topology were determined from a combination of H^{α}, C^{α}, C^{β}, and C′ secondary chemical shift analysis and identification of unambiguous NOEs. Secondary chemical shifts were calculated using published random coil values for H^{α}, C^{α}, C^{β}, and C′ nuclei (*30*). ^{1}H, ^{15}N, and ^{13}C resonances were externally referenced relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) (0 ppm) using the ratios 0.251 449 530 and 0.101 329 188 for ^{13}C and ^{15}N, respectively (*31*). Appropriate corrections to the random coil shifts were made for residues preceding a proline residue (*32*). Heteronuclear {^{1}H}-^{15}N steady-state NOE experiments (*33*) were recorded with a recycle delay of 6 s for the reference spectrum, while for the spectrum recorded with presaturation, a 3 s recycle delay followed by a 3 s period of ^{1}H saturation was employed.

Distance Restraints. Interproton distance restraints were determined from cross-peak intensities. Distance restraints of 2.7, 3.3, 5.0, 6.0, and 6.5 Å were employed for cross-peaks of strong, medium, medium-weak, weak, and very weak intensity, respectively, in the shared time 3D ^{15}N/^{13}C-NOESY-HSQC (*28*), ^{15}N-NOESY-HSQC (*22*), 2D homonuclear NOESY (*34*), and ^{13}C-HMQC-NOESY (*29*) spectra. An additional 0.5 Å was added to the bounds for NOEs involving methyl protons (*35*), and standard pseudoatom corrections for aromatic protons and for residues with degenerate geminal methylene proton resonances were applied. Hydrogen bond restraints were derived for residues in the β sheet from a combination of NOE and amide proton exchange data.

Torsion Angle Restraints. Torsion angle restraints were determined on the basis of $^{3}J_{\mathrm{HNH}\alpha}$ coupling constants from the HNHA (*36*, *37*) spectrum. Backbone ϕ dihedral angle restraints were applied to 58 residues. Residues with $^{3}J_{\mathrm{HN,H}\alpha} < 4.0$ Hz were constrained to $-120° < \phi < -20°$, and residues with $^{3}J_{\mathrm{HN,H}\alpha} > 8.0$ Hz were restrained to $-170° < \phi < -70°$ in both DYANA and AMBER structure calculations. In addition, ϕ/ψ dihedral angle restraints were added for residues identified as occupying α or β regions of the Ramachandran plot based on their C^{α}, C^{β}, H^{α}, and C′ chemical shifts. Residues identified in the α region were restrained to $-120° < \phi < -20°$ and $-80° < \psi < 20°$, and residues defined in the β region were restrained to $-170° < \phi < -70°$ and $60° < \psi < 180°$. Nonglycine residues without experimentally determined ϕ restraints were restrained to negative ϕ angles, unless secondary C^{β} shifts suggested the possibility of a positive ϕ angle.

Stereospecific assignments of $C^{\beta}H_2$ resonances were made by qualitative assessment of $^{3}J_{\mathrm{H}\alpha,\mathrm{H}\beta}$ and $^{3}J_{\mathrm{N,H}\beta}$ couplings, with confirmation from inter and intraresidue NOEs. Stereospecific $C^{\beta}H_2$ assignments were made for 21 of the Asp, Asn, Gln, His, Leu, Ser, Tyr, and Phe residues, and χ_1 angles were determined from analysis of the HNHB (*38*), DQF-COSY (*25*), and HACAHB-COSY (*39*) spectra and restrained to one of the three staggered rotameric conformations $\chi_1 = -60\ (\pm 40)°$, $60\ (\pm 40)°$, or $180\ (\pm 40)°$. Stereospecific assignments for 4 Val γ-methyl and 4 Leu δ-methyl groups which were not obscured by overlap were determined from the NOESY data. The ω angles of the peptide backbone were constrained to $180\ (\pm 10)°$ for all residues in the AMBER calculations. Stereospecific assignments for several of Val residues were obtained by a combination of NOESY data and ^{13}C-{^{13}CO} (*40*) and ^{13}C-{^{15}N} (*41*) spin−echo difference CT-HSQC spectra.

Structure Calculations. Structures were generated from random atom coordinates using DYANA (*42*) and were further refined with steps of torsion angle dynamics and energy minimization using AMBER 6 and 7 (*43*). An initial 1700 randomized starting structures were generated. A total of 701 experimental restraints were used for the initial DYANA calculations. The structures were refined with 4000 steps of torsion angle dynamics, ending with 1000 steps of energy minimization in DYANA. Of the 1700 structures, 656 structures with target functions less than 0.2 kcal/mol were selected for further refinement in a vacuum using AMBER 6.0 (*43*). Structures were refined using a total of 1141 unambiguous interproton distance restraints. The net charges on the charged side chains and the N and C termini were reduced to 20% of their normal value. Structures were subjected to 2000 steps of energy minimization followed by four cycles of vacuum simulated annealing from 1000 to 6 K, 20 ps each, followed by a final 2000 steps of energy minimization. The vacuum simulated annealing involved an initial heating of the structures to 1000 K over 2 ps, followed by 4 ps of molecular dynamics calculations, and a final simulated annealing protocol where the temperature was reduced to 6 K during 12 ps. The temperature was lowered with τ held constant at 2.0 for 5 ps, then at 1.0 for 5 ps, 0.5 for 2 ps, and finally at 0.05 for the remaining 2 ps. Each round was of a duration of 20 ps, consisting of 20 000 steps of 1 fs each. Experimentally determined distance and torsion angle restraints were applied with force constants of 50 kcal mol^{-1} Å^{-1} and 150 kcal mol^{-1} rad^{-2}, respectively. The force constant associated with the chirality restraints (ω dihedral angles) was 150 kcal mol^{-1} rad^{-2}. The experimental restraints were introduced into the calculations progressively, to ensure maximum sampling of conformational space, by increasing their weighting from an initial value of 0.1 to 1.0 (normal level) over the first 14 ps of the AMBER calculations.

The 200 structures with the lowest restraint and AMBER energies were subjected to further AMBER refinement cycles. Finally, a family of 92 structures, selected on the basis of restraint violation energies and AMBER energies, was refined in AMBER 7 using the generalized Born (GB) continuum solvent model (*44*, *45*). The AMBER 7 calculations involved an initial heating of the structures to 1000 K over 4 ps, followed by 4 ps of molecular dynamics calculations, and a final simulated annealing protocol where the temperature was reduced to 0 K during 12 ps. The

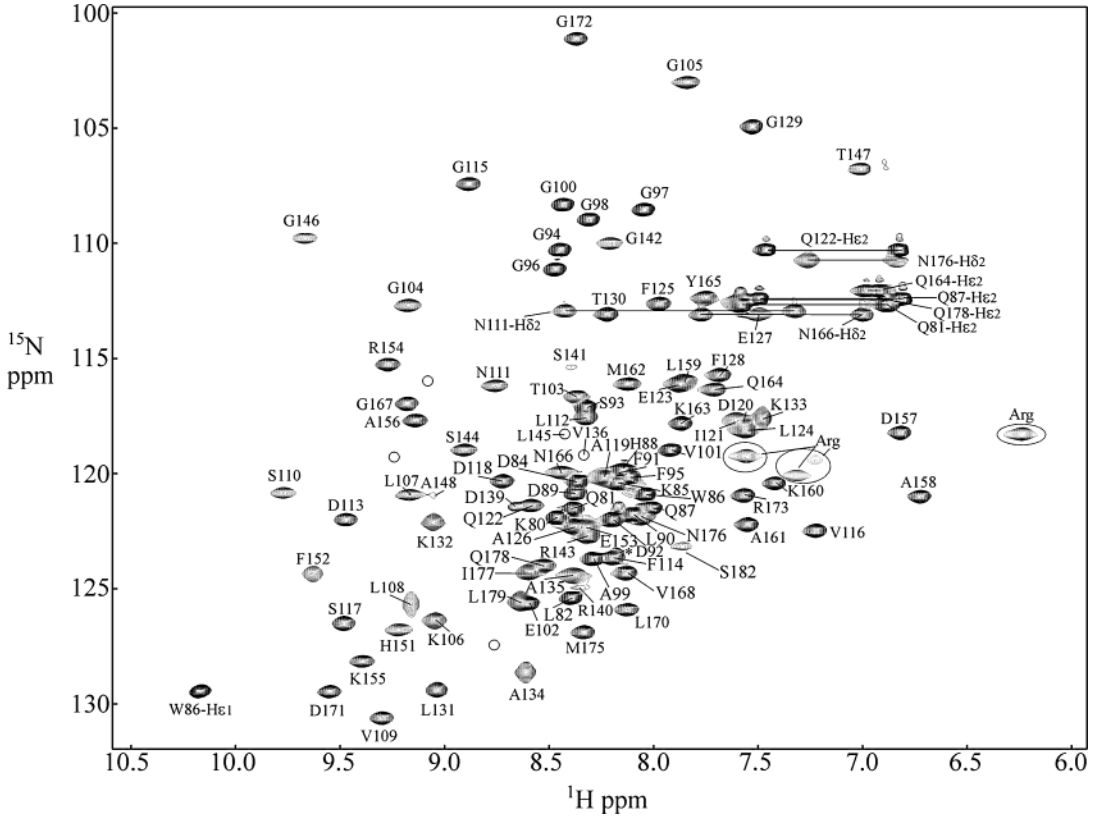

FIGURE 2: ^{1}H-^{15}N HSQC spectrum of ALY77−182 recorded at 800 MHz. The * corresponds to Leu82 preciding cis-Pro83.

experimental restraints were introduced into the calculations progressively by increasing their weighting from an initial value of 0.0−1.0 over the first 12 ps of the calculations. A final family of 32 structures was selected on the basis of low restraint and AMBER energies. The quality of these structures was determined using PROCHECK-NMR (*46*) and PROMOTIF-NMR (*47*).

Coordinates and Assignments. Resonance assignments have been deposited in the BioMagResBank (accession number 5764). The coordinates for the 32 structures have been deposited in the Protein Data Bank (accession number 1NO8).

RESULTS AND DISCUSSION

NMR Resonance Assignments of ALY77−182. The boundaries for the ALY77−182 construct were determined from previous domain mapping experiments (*1*), secondary structure prediction, and analysis of the expression and solubility of different constructs with N-terminal and C-terminal extensions. A construct composed of residues 77−182 of mouse ALY (ALY77−182) exhibited the most favorable solution properties. ALY77−182 was soluble to 0.5 mM under high salt conditions (100 mM NaCl).

^{1}H-^{15}N-correlated NMR spectrum of ALY77−182 exhibits signals characteristic of a folded protein plus sharp overlapped resonances indicative of flexible regions (Figure 2). Extensive backbone resonance assignments (93% of residues) were made using sequential connectivities observed in the ^{15}N edited NOESY and triple resonance data (*48*, *49*). The backbone amide proton resonances of residues Arg78 and His137 were missing or broadened due to amide proton exchange. Although chemical shift dispersion of the majority of the backbone H^N protons is good, overlap was observed

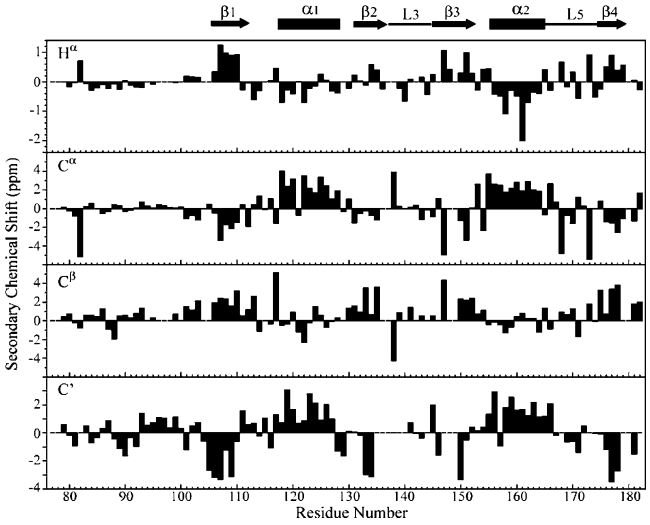

FIGURE 3: Secondary chemical shifts plotted for H^{α}, C^{α}, C^{β}, and C′ nuclei of the ALY77−182 protein. The cartoon on top of the figure depicts the location of the β strands and α helices.

for many of the loop regions as well as the N-terminal flexible region. Data were collected at two temperatures, 23and 16 °C, and two pHs, 5.9 and 6.5, to assign resonances that were in the NMR intermediate time scale, were degenerate or in rapid exchange with the solvent. Side chain assignments were made using triple resonance spectra. Most of the aromatic ring protons were identified and assigned from the double-quantum filtered COSY spectrum.

Elements of secondary structure were identified on the basis of the secondary chemical shifts of the H^{α}, C^{α}, C^{β}, and C′ nuclei (Figure 3) (*30*). This information combined with NOESY data allowed the structure of the central β sheet to be defined (Figure 4).

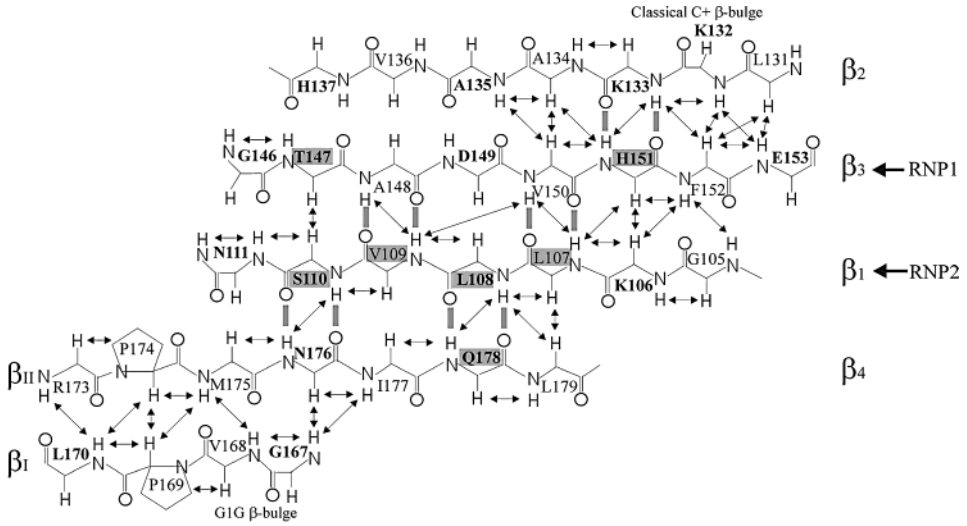

FIGURE 4: Summary of some of the backbone NOE connectivities that define the topology of the β sheet in ALY77−182. The schematic also includes residues that form the small β sheet (β_I−β_{II}) in loop L_5. Residues surrounded by rectangles indicate slow H^N exchange (present after 1 h of resuspension in D_2O, at room temperature). Side chains of residues in bold face the surface.

Table 1: NMR Restraints and Structural Statistics for 32 Refined Structures

A. NMR Restraints	
total NMR distance restraints	1141
distance restraints	1121
intraresidue (i,i)	492
sequential ($\|i - j\| = 1$)	292
medium-range ($1 \leq \|i - j\| \leq 4$)	144
long-range ($\|i - j\| > 4$)	193
hydrogen bonds	10
total dihedral angle restraints	91
ϕ	58
ψ	14
χ_1	19
B. Violation Statistics	
number of NOE violations > 0.09 Å	1
maximum NOE violation (Å)	0.12
number of torsion angle restraints > 2.5°	0
energies (kcal/mol)	
mean restraint violation energy	0.93 ± 0.22
mean amber energy	−4319.08 ± 13.75
C. Average RMS Deviations from Ideal Covalent Geometry	
bond lengths (Å)	0.0101 ± 0.0001
bond angles (deg)	2.09 ± 0.03

D. PROCHECK Statistics for ϕ and ψ (G105−S182)

	GB	vacuum
residues in most favored regions (%)	88.8	72.7
residues in additional allowed regions (%)	10.2	24.4
residues in generously allowed regions (%)	0.8	1.3
residues in disallowed regions (%)	0.2	1.5

Tertiary Structure Determination. A total of 1141 experimental restraints were used for structure calculations (Table 1). Hydrogen bonds were included for slowly exchanging amides in stretches of regular secondary structure (β sheet only) where hydrogen-bond acceptors were unambiguous. Of the structures calculated, 32 were picked which had converged to the lowest energies. The final family of 32 structures had no NOE violations greater than 0.12 Å and no dihedral angle violations greater than 2.5°. A superposition of the 32 structures is shown in Figure 5a. The backbone atoms involved in secondary structural elements are well defined in this family, with low RMSD from the mean structure.

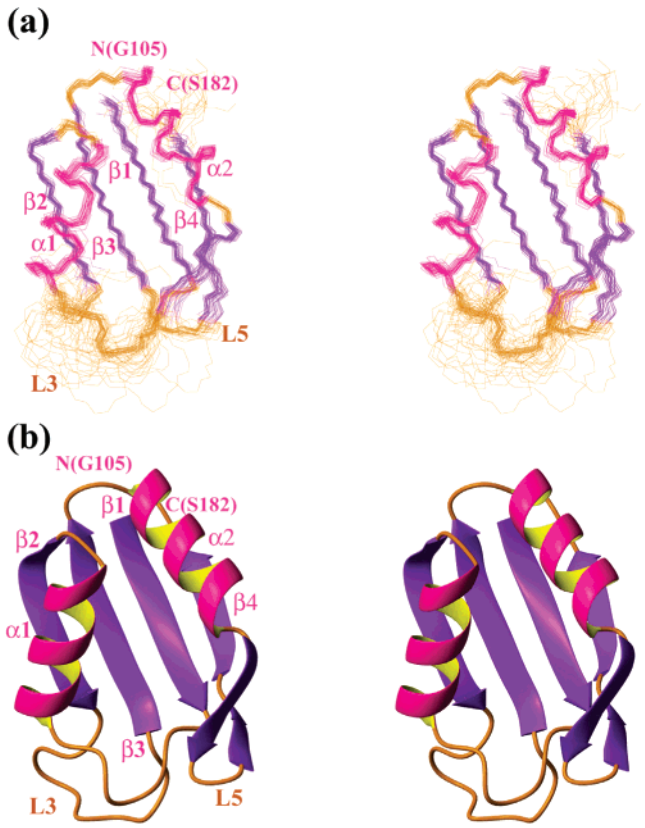

FIGURE 5: (a) Stereoview of the backbone (C^α, C', N, O) superposition of the 32 structures of the ensemble with β strands in purple, α helices in magenta and loops in gold. Only residues Gly105−Ser182 are shown. (b) Stereoview showing a ribbon diagram of a C^α trace of the lowest energy structure of ALY77−182. Only residues Gly105 to Ser182 are shown. These figures were generated in MOLMOL (*79*).

The structures were generated from random atom coordinates using DYANA (*42*) and refined with DYANA and the program AMBER 6.0 in a vacuum (*43*). Structures were selected based on the lowest restraint and AMBER energies and submitted to several molecular dynamics cycles using a generalized Born continuum solvent model. The use of GB potentials has been shown to improve the accuracy of

Table 2: Difference in RMSDs between Vacuum Calculations and GB

	vacuum (RMSDs pairwise/RMSDs from mean)	GB (RMSDs pairwise/RMSDs from mean)
Residues G105−S182 (Å)		
bb (N, C^{α}, C′)	2.35 ± 0.58/1.64 ± 0.38	2.01 ± 0.54/1.40 ± 0.40
heavy (N, C, O, S)	3.06 ± 0.58/2.14 ± 0.35	2.67 ± 0.52/1.86 ± 0.39
Helix 1 and 2 (S117−F128, R154−Y165) (Å)		
bb (N, C^{α}, C′)	0.74 ± 0.23/0.52 ± 0.18	0.47 ± 0.16/0.32 ± 0.12
heavy (N, C, O, S)	1.68 ± 0.27/1.17 ± 0.18	1.39 ± 0.20/0.97 ± 0.11
β-Sheet (G105−S110, L131−V136, T147−F152, N176−L179) (Å)		
bb (N, C^{α}, C′)	0.88 ± 0.17/0.61 ± 0.12	0.56 ± 0.14/0.39 ± 0.10
heavy (N, C, O, S)	1.55 ± 0.21/1.08 ± 0.13	1.22 ± 0.22/0.85 ± 0.17

calculated solution structures, giving similar results to explicit solvent calculations (*50*). The restraint and angle violations for the final family of structures improved after using the generalized Born protocol. Structures from refinement in a vacuum had a mean restraint violation energy of 4.0 kcal/mol, whereas the structures after calculations with the GB potential had a mean restraint violation energy of 0.93 kcal/mol. Even more dramatic was the improvement in the quality of structures assessed by PROCHECK-NMR (*46*) (Table 2). Statistics for residues in the ϕ and ψ regions for residues Gly105-Ser182 in the Ramachandran plot improved from 72.7% to 88.6% in the most favored regions (Table 1). RMSDs for different regions of the structures improved significantly (Table 2).

Tertiary Structure of ALY77−182. The structure of the conserved RBD domain in ALY consists of four β strands β_1(Gly105−Ser110), β_2(Leu131−Val136), β_3(Thr147−Phe152), and β_4(Asn176−Leu179) forming an antiparallel β sheet. Two alpha helices α_1(Ser117-Phe128) and α_2-(Arg154-Tyr165), almost perpendicular to each other, pack against one face of the β sheet (Figure 5b), making hydrophobic contacts via the side chains facing inward into the β sheet. The exposed face of the β sheet (Figure 6a) presents to the solvent hydrophobic and charged side chains that are potential binding sites for interaction with RNA and other proteins (Figure 6b,c).

Two long loops facing the same side of the main fold, L_3(His137−Gly146) and L_5(Asn166−Met175), have different relaxation properties. The resonances of residues in loop L_3 are extremely broad, whereas those of residues in loop L_5 are sharp. Due to the broadness of the resonances, the conformation of loop L_3 could not be determined unambiguously. The $\{^1H\}$-^{15}N heteronuclear NOEs for residues in loop L_3 and L_5 indicated that those loops are not fully flexible (Figure 7). Loop L_5 consists of 10 residues that form a small antiparallel β sheet between Gly167−Leu170 (β_I) and Arg173−Met175 (β_{II}). Residues Asp171 and Gly172 form a type I′ β turn (*51*, *52*). Residues Gly167, Val168, and Met175 form a G1G type β bulge (*53*).

The N-terminal tail consisting of residues Ser77 to Gly104 is highly disordered, as deduced from the secondary chemical shifts, the lack of medium-range and long-range NOEs and the $\{^1H\}$-^{15}N heteronuclear NOEs (Figures 3 and 7). Nevertheless, residues Asp84 to Phe95 exhibit some restriction of motion, as indicated by $\{^1H\}$-^{15}N heteronuclear NOEs in the range of 0.1−0.2. Secondary chemical shifts from Ser93 to Gly100 are consistent with sampling in the helical region of ϕ,ψ space (Figure 3).

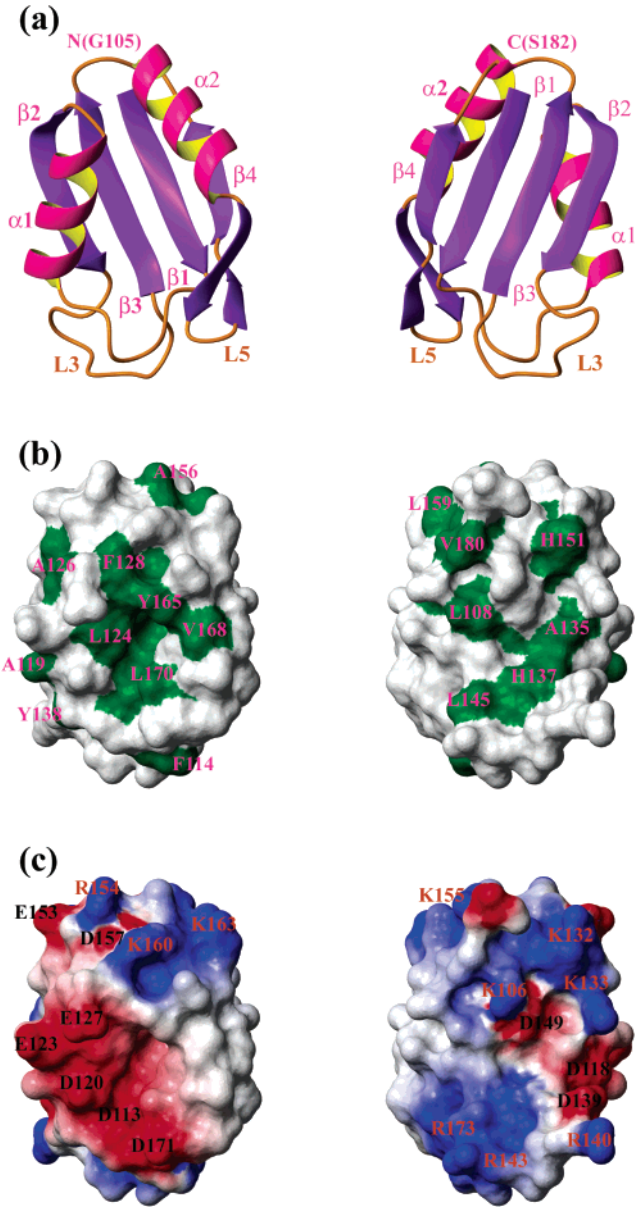

FIGURE 6: (a) Ribbon diagram denoting the surface orientations with respect to the Cα trace. Molecular contact surfaces of a low restraint violation structure generated with MOLMOL (*79*). (b) Surface representation denoting in green the residues with hydrophobic side chains in the surface of ALY77−182 and (c) Surface representation showing the charge potential at the surface of ALY77−182.

Comparison with Other RBD Structures. Proteins that contain RBD domains can have single or multiple RBDs with

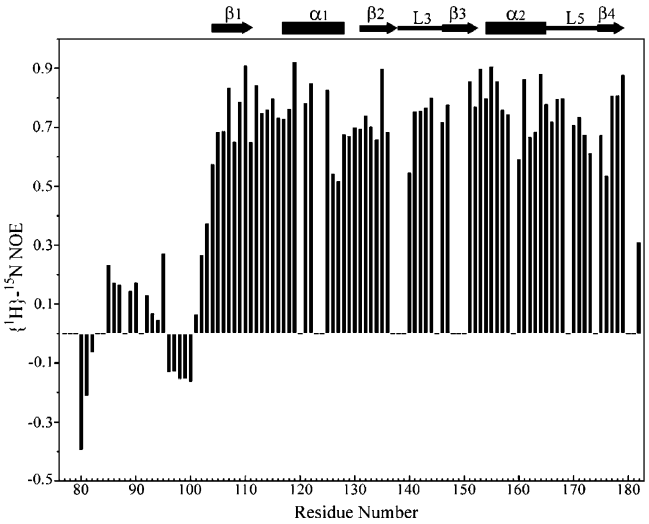

FIGURE 7: {^{1}H}-^{15}N heteronuclear NOE plotted as a function of residue number. The secondary structural elements are shown schematically.

different types of recognition. In the case of ALY, the folding topology for residues Gly105 to Val180 is similar to that in the structures of other well characterized RBDs that bind RNA (Figure 5a,b) (*54*−*70*).

Strands β_1 and β_3 in ALY77−182 contain the conserved RNP2 and RNP1 submotifs, respectively, indicated in Figure 1. A particular characteristic of the RNP submotifs in the β sheet is partially conserved Phe/Tyr and well-conserved Phe/Tyr residues at positions equivalent to Leu108 and Asp149 in ALY, respectively (Figure 1). In snRNP U1A RBD1 and U2B″ RBD1, these aromatic residues make contacts with RNA and with each other. The absence of the highly conserved Phe/Tyr in the ALY RNP1 submotif is characteristic of ALY and its homologues and may indicate a unique mechanism of RNA binding by the RBD in ALY like proteins.

β bulges within the β sheet have been found to contribute to the specificity of protein and RNA recognition. β bulges have been observed, for example, in β strands β_1 and β_4 in hnRNP C (*54*), β_2 of Sxl RBD2 (*62*), and snRNP U1A RBD1 (*55*) but are not present in hnRNP A1 (*67*−*69*). In ALY, a classic β bulge is formed by Lys132 and Lys133 instead of bulky hydrophobic and acidic residues as in the case of other RBDs such as Sxl RBD2 and snRNP U1A RBD1 (Figure 1). The presence of two positively charged side chains at these positions is characteristic among close homologues of ALY such as DIP1, DIP2, the ALY yeast homologue-YRA1, and the yeast protein MLO3 (Figure 1), which have unique functions.

Residues in the loop L_3 and on the surface of the β sheet are poorly conserved among RBDs of known structure (Figure 1). ALY77−182 contains nine residues between β_2 and β_3, compared to five residues in U1A RBD2, six residues in nucleolin RBD2, seven residues in U1A RBD1 and U2B″ RBD1, eight residues in PABP RBD2 and nucleolin RBD1, 10 residues in hnRNP A1 RBD1 and RBD2, Sxl RBD1 and RBD2, HUD RBD1 and RBD2, and PABP RBD1, and a two-residue β-hairpin in hnRNP C1. Amide proton resonance signals for residues in this loop are extremely broad for ALY77−182.

In known RBD structures, helices α_1 and α_2 are of similar length, but the interhelix crossing angles vary from 90° to 120°. The helices in the RBD2 of Sxl (93°), hnRNP C (90°) (*54*), hnRNP A1 RBD1 and RBD2 (97°, 94°), and nucleolin RBD2 (98°) are almost perpendicular with respect to each other. In the case of ALY77−182, the helices α_1 and α_2 have a crossing angle of 111 degrees, similar to that of the RNA binding RBDs of U1A (106°) and nucleolin (111°). It is interesting to note that the RBDs in Sxl and U1A that do not apparently participate in RNA interactions have larger crossing angles (120° and 119°, respectively).

Loop L_5, which connects the second helix, α_2, and the C-terminal β strand, β_4, exhibits low conservation among RBDs (Figure 1). This loop in U2B″ RBD1 (*63*), hnRNP A1 RBD2 (*69*), U1A RBD1, and Sxl RBD2 (*61*) makes few contacts with nucleic acids. Loop L_5 in ALY contains close sequence and length similarities to loop L_5 in U2B″ RBD1, U1A RBD1, PAB1 RBD2, and nucleolin RBD1 and RBD2. Gly167, Pro169, Pro174, and Met175 in ALY are identical in U2B″ RBD1 and U1A RBD1 (Figure 1). Also, residues Val168 and Leu170 are conserved hydrophobic in ALY, and a conserved positively charged residue that makes contacts with RNA in U1A RBD1 and U2B″ RBD1 is at the conserved position of Arg173.

Even though ALY77−182 is similar in fold to other RBDs from other proteins, the surface residues of ALY RBD are different from those conserved in other RBDs. These differences confer specificity in molecular recognition and may explain why ALY has different predilections for interactions. The backbone resonances of the residues within the RNP2 (β_1) and RNP1 (β_3) submotifs in ALY77−182 that are exposed to the surface (Asp149, His151, Leu108, Ser110), at positions that are involved in RNA interactions in other RBDs, are extremely broad, suggesting residual mobility at the backbone level. Line broadening propagates to residues within the submotifs whose side chains form part of the hydrophobic core (Ala148, Val150, and Leu107).

Implications for the Function of ALY. The main region in ALY77−182 that could facilitate RNA interactions is composed of the exposed face of the β sheet and loop L_3. The opposite face exposes side chains in loop L_5 that could modulate hydrophobic interactions with another molecular surface. It should also be noted that the N-terminal region of ALY77−182 which corresponds to residues 77−104 is flexible and without regular structure. This N-terminal "tail" contains several aliphatic and aromatic side chains that could also modulate interactions.

ALY77−182 possesses the conserved RNP1 and RNP2 submotifs that are characteristic of RBDs. However, a highly conserved Phe/Tyr in the middle of the RNP1 submotif, which participates in base stacking interactions in known RBD:RNA complexes, is replaced by an aspartate (Asp149) in ALY. The absence of Phe/Tyr at this position appears unique to ALY and its homologues and must account for their unique functions. This highly conserved aromatic residue is substituted by polar residues in a subset of RBD sequences. The *Arabidopsis* DIP1 and DIP2 proteins, which have Glu at this position, interact with the DNA binding domain of plant poly(ADP-ribose) polymerase (*71*). MLO3 from *Schizosaccharomyces pombe* (Thr substitution) appears to play a role in chromosome segregation in mitosis (*72*). YRA1 (Asn substitution) is required for mRNA export in

yeast, interacts with poly(A)$^+$ RNA, has in vitro RNA−RNA annealing activity, and binds to Mex67p which associates in vivo with nuclear pores and poly(A)$^+$ RNA (*4*).

ALY has been determined to bind to the mRNA export factor TAP (*2*). This interaction is conserved in the ALY yeast homologue YRA1, which interacts with the TAP yeast homologue Mex67p (*4*). The structure of the minimal region of TAP required for interaction with ALY contains two domains: a noncanonical RBD and a conserved leucine rich region (LRR) domain similar to that of U2A′ (*73*). U2A′ interacts with the U2B″ RBD1, which binds to U2 snRNA only in the presence of U2A′ (*63*). In the case of TAP, it interacts on its own with the cis-acting constitutive transport element (CTE) of simian type D retroviruses (*74*), suggesting that the LRR and RBD regions of TAP might come into contact to bind to the CTE in a manner similar to the U2B″−U2A′−U2 snRNA interactions (*73*). Since the presence of ALY is required for TAP binding to cellular mRNAs and considering that the RBD in ALY has more structural and sequence similarities to U2B″ than the non canonical RBD of TAP, it is possible that association of ALY with the LRR of TAP is necessary for interaction with cellular mRNAs. On the other hand, on the basis of the mode of RNA interaction by other RBDs such as RBD1 and RBD2 of the polyA binding protein 1 (*64*), RBD1 and RBD2 of Sex Lethal (*61*, *62*), RBD1 and RBD2 of HuD (*70*), and RBD1 and RBD2 of nucleolin (*65*), it is also possible that the concerted interactions of ALY and the noncanonical RBD of TAP are necessary for complex formation. The regions of ALY responsible for interaction with TAP and mRNA were previously determined as the variable regions flanking the RBD domain in ALY (*75*). The variable N and C terminal regions in ALY, however, exhibit low sequence conservation and are composed of repeats of positively charged residues that could lead to nonspecific interactions with TAP and mRNA. It is also possible that the RBD domain of ALY is directly involved in the interaction with TAP. Low-affinity interactions between the RBD of ALY and TAP were observed in yeast two hybrid experiments (M. Andersen-Landes, unpublished results, 2002). The interaction may be strengthened in the ternary complex.

The N-terminal helix in ALY77−182 (α_1) exposes negatively charged side chains consisting of D118, D120, E123, and E127 (Figure 6c). Clusters of negatively charged residues at the α_1 position seem common among other RBDs (Figure 1). Interestingly, DIP1 and DIP2 contain a putative modification domain at the same position. It has been suggested that this cluster of Glu residues in DIP1 and DIP2 might undergo poly(ADP-ribosyl)ation by the plant homologue of the Zn-finger poly(ADP-ribose) polymerase (PARP1) modulating their interaction (*71*). ALY has been found to interact with DNA binding transcription factors as well as RNA binding proteins. These interactions could be modulated via the exposed negative surface of α_1 in the RBD in ALY and the nucleic acid binding surfaces of the DNA and RNA binding proteins that have been found to interact with ALY. One of these interactions is the interaction of ALY with the DEAD-box helicase UAP56. Studies have shown that UAP56 couples ALY with transcription as well as mRNA export (*13*, *14*).

In summary, ALY is an important factor for mRNA export and also for enhancing transcriptional activation by some nuclear factors. The three-dimensional structure of ALY77−182 and its comparison with RBD proteins of known structure provides insights into the differences in ALY that could account for its unique function. Comparisons to its homologues and other RBD proteins with close sequence homology has allowed us to make predictions regarding its mechanism of interaction based on its three-dimensional structure.

ACKNOWLEDGMENT

We gratefully acknowledge Dr. Brian M. Lee, Dr. Micah Gearhart, Mindy Andersen-Landes, Theodore Foss, Linda Tennant, Leonard Kaljevic, Dr. John Chung, and Dr. Gerard Kroon for assistance. Dr. David Case provided valuable guidance in the implementation of the GB potentials. We thank Drs. Sonja Dames, Roberto De Guzman, Annette Atkins, and Jaime Pascual for valuable discussions.

REFERENCES

1. Bruhn, L., Munnerlyn, A., and Grosschedl, R. (1997) ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRα enhancer function, *Genes Dev. 11*, 640−653.
2. Stutz, F., Bachi, A., Doerks, T., Braun, I. C., Seraphin, B., Wilm, M., Bork, P., and Izaurralde, E. (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export, *RNA 6*, 638−650.
3. Virbasius, C. M., Wagner, S., and Green, M. R. (1999) A human nuclear-localized chaperone that regulates dimerization, DNA binding, and transcriptional activity of bZIP proteins, *Mol. Cell 4*, 219−228.
4. Strässer, K., and Hurt, E. (2000) Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export, *EMBO J. 19*, 410−420.
5. Reed, R., and Magni, K. (2001) A new view of mRNA export: separating the wheat from the chaff, *Nat. Cell Biol. 3*, E201−E204.
6. Zhou, Z., Luo, M. J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000) The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans, *Nature 407*, 401−405.
7. Reichert, V. L., Le Hir, H., Jurica, M. S., and Moore, M. J. (2002) 5′ exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly, *Genes Dev. 16*, 2778−2791.
8. Le Hir, H., Izaurralde, E., Maquat, L. E., and Moore, M. J. (2000) The spliceosome deposits multiple proteins 20−24 nucleotides upstream of mRNA exon-exon junctions, *EMBO J. 19*, 6860−6869.
9. Kataoka, N., Diem, M. D., Kim, V. N., Yong, J., and Dreyfuss, G. (2001) Magoh, a human homologue of *Drosophila* mago nashi protein, is a component of the splicing-dependent exon-exon junction complex, *EMBO J. 20*, 6424−6433.
10. Lykke-Andersen, J., Shu, M. D., and Steitz, J. A. (2001) Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1, *Science 293*, 1836−1839.
11. Kim, V. N., Kataoka, N., and Dreyfuss, G. (2001) Role of the nonsense-mediated decay factor hUpf3 in the splicing- dependent exon-exon junction complex, *Science 293*, 1832−1836.
12. Kim, V. N., Yong, J., Kataoka, N., Abel, L., Diem, M. D., and Dreyfuss, G. (2001) The Y14 protein communicates to the cytoplasm the position of exon-exon junctions, *EMBO J. 20*, 2062−2068.
13. Luo, M. L., Zhou, Z., Magni, K., Christoforides, C., Rappsilber, J., Mann, M., and Reed, R. (2001) Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly, *Nature 413*, 644−647.
14. Strässer, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondon, A. G., Aguilera, A., Struhl, K., Reed, R., and Hurt, E. (2002) TREX is a conserved complex coupling transcription with messenger RNA export, *Nature 417*, 304−308.

15. Delaglio, F., Grzesiek, S., Vuister, G. W., Guang, Z., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
16. Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603−614.
17. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein, *J. Magn. Reson. 96*, 432−440.
18. Wittekind, M., and Mueller, L. (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins, *J. Magn. Reson. 101*, 201−205.
19. Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J. Am. Chem. Soc. 114*, 6291−6293.
20. Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation, *J. Magn. Reson., Ser. A 109*, 129−133.
21. Grzesiek, S., Anglister, J., and Bax, A. (1993) Correlation of backbone amide and aliphatic side chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of ^{13}C magnetization, *J. Magn. Reson., Ser. B 101*, 114−119.
22. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of ^1H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^1H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to Interleukin 1β, *Biochemistry 28*, 6150−6156.
23. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) ^1H-^1H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C-enriched proteins, *J. Magn. Reson. 88*, 425−431.
24. Ikura, M., Kay, L. E., and Bax, A. (1991) Improved three-dimensional ^1H-^{13}C-^1H correlation spectroscopy of a ^{13}C-labeled protein using constant-time evolution, *J. Biomol. NMR 1*, 299−304.
25. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering, *Biochem. Biophys. Res. Commun. 117*, 479−485.
26. Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy, *J. Magn. Reson. 53*, 521−528.
27. Messerle, B. A., Wider, G., Otting, G., Weber, C., and Wüthrich, K. (1989) Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with H_2O solutions, *J. Magn. Reson. 85*, 608−613.
28. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1994) Simultaneous acquisition of ^{15}N- and ^{13}C-edited NOE spectra of proteins dissolved in H_2O, *J. Magn. Reson., Ser. B 103*, 197−201.
29. Fesik, S. W., and Zuiderweg, E. R. P. (1988) Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra, *J. Magn. Reson. 78*, 588−593.
30. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination, *Methods Enzymol. 239*, 363−392.
31. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^1H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR 6*, 135−140.
32. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^1H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids: I. Investigations of nearest neighbor effects, *J. Biomol. NMR 5*, 67−81.
33. Grzesiek, S., and Bax, A. (1993) The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements, *J. Am. Chem. Soc. 115*, 12593−12594.
34. Macura, S., and Ernst, R. R. (1980) Elucidation of cross-relaxation in liquids by two-dimensional N.M.R. spectroscopy, *Mol. Phys. 41*, 95−117.
35. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance, distance geometry, and restrained molecular dynamics, *Biochemistry 26*, 8012−8023.
36. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: A new approach for measuring homonuclear three-bond $J(H^NH^\alpha)$ coupling constants in ^{15}N-enriched proteins, *J. Am. Chem. Soc. 115*, 7772−7777.
37. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) Measurement of H^N-H^α J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods, *J. Biomol. NMR 4*, 871−878.
38. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side chain Hβ resonances in larger proteins, *J. Magn. Reson. 95*, 636−641.
39. Grzesiek, S., Kuboniwa, H., Hinck, A. P., and Bax, A. (1995) Multiple-quantum line narrowing for measurement of Hα-Hβ J coupling in isotopically enriched proteins, *J. Am. Chem. Soc. 117*, 5312−5315.
40. Grzesiek, S., Vuister, G. W., and Bax, A. (1993) A simple and sensitive experiment for measurement of J_{CC} couplings between backbone carbonyl and methyl carbons in isotopically enriched proteins, *J. Biomol. NMR 3*, 487−493.
41. Vuister, G. W., Wang, A. C., and Bax, A. (1993) Measurement of three-bond nitrogen−carbon *J* couplings in proteins uniformly enriched in ^{15}N and ^{13}C, *J. Am. Chem. Soc. 115*, 5334−5335.
42. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
43. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1999) in *AMBER 6*, University of California, San Francisco.
44. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics, *J. Am. Chem. Soc. 112*, 6127−6129.
45. Bashford, D., and Case, D. A. (2000) Generalized born models of macromolecular solvation effects, *Annu. Rev. Phys. Chem. 51*, 129−152.
46. 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*, 477−486.
47. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF-A program to identify and analyze structural motifs in proteins, *Protein Sci. 5*, 212−220.
48. Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
49. Grzesiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins, *J. Biomol. NMR 3*, 185−204.
50. Xia, B., Tsui, V., Case, D. A., Dyson, H. J., and Wright, P. E. (2002) Comparison of solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model and with explicit water, *J. Biomol. NMR 22*, 317−331.
51. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) Conformation of β-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering, *J. Mol. Biol. 206*, 759−777.
52. Guruprasad, K., and Rajkumar, S. (2000) Beta-and gamma-turns in proteins revisited: a new set of amino acid turn-type dependent positional preferences and potentials, *J. Biosci. 25*, 143−156.
53. Chan, A. W., Hutchinson, E. G., Harris, D., and Thornton, J. M. (1993) Identification, classification, and analysis of beta-bulges in proteins, *Protein Sci. 2*, 1574−1590.
54. Wittekind, M., Gorlach, M., Friedrichs, M., Dreyfuss, G., and Mueller, L. (1992) ^1H, ^{13}C and ^{15}N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins, *Biochemistry 31*, 6254−6265.
55. Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A, *Nature 348*, 515−520.
56. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin, *Nature 372*, 432−438.
57. Avis, J. M., Allain, F. H. T., Howe, P. W. A., Varani, G., Nagai, K., and Neuhaus, D. (1996) Solution structure of the N-terminal

RNP domain of U1A protein: The role of C-terminal residues in structure stability and RNA binding, *J. Mol. Biol. 257*, 398−411.
58. Allain, F. H. T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation, *Nature 380*, 646−650.
59. Lu, J. R., and Hall, K. B. (1997) Tertiary structure of RBD2 and backbone dynamics of RBD1 and RBD2 of the human U1A protein determined by NMR spectroscopy, *Biochemistry 36*, 10393−10405.
60. Inoue, M., Muto, Y., Sakamoto, H., Kigawa, T., Takio, K., Shimura, Y., and Yokoyama, S. (1997) A characteristic arrangement of aromatic amino acid residues in the solution structure of the amino-terminal RNA-binding domain of *Drosophila* Sex-lethal, *J. Mol. Biol. 272*, 82−94.
61. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) Structural basis for recognition of the *tra* mRNA precursor by the Sex- lethal protein, *Nature 398*, 579−585.
62. Lee, A. L., Kanaar, R., Rio, D. C., and Wemmer, D. E. (1994) Resonance assignments and solution structure of the second RNA-binding domain of sex-lethal determined by multidimensional heteronuclear magnetic resonance, *Biochemistry 33*, 13775−13786.
63. Price, S. R., Evans, P. R., and Nagai, K. (1998) Crystal structure of the spliceosomal U2B″-U2A′ protein complex bound to a fragment of U2 small nuclear RNA, *Nature 394*, 645−650.
64. Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein, *Cell 98*, 835−845.
65. Allain, F. H. T., Gilbert, D. E., Bouvet, P., and Feigon, J. (2000) Solution structure of the two N-terminal RNA-binding domains of nucleolin and NMR study of the interaction with its RNA target, *J. Mol. Biol. 303*, 227−241.
66. Allain, F. H. T., Bouvet, P., Dieckmann, T., and Feigon, J. (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin, *EMBO J. 19*, 6870−6881.
67. Garrett, D. S., Lodi, P. J., Shamoo, Y., Williams, K. R., Clore, G. M., and Gronenborn, A. M. (1994) Determination of the secondary structure and folding topology of an RNA binding domain of mammalian hnRNP A1 protein using three-dimensional heteronuclear magnetic resonance spectroscopy, *Biochemistry 33*, 2852−2858.
68. Shamoo, Y., Krueger, U., Rice, L. M., Williams, K. R., and Steitz, T. A. (1997) Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 Å resolution, *Nat. Struct. Biol. 4*, 215−222.
69. Ding, J. Z., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R., and Xu, R. M. (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA, *Genes Dev. 13*, 1102−1115.
70. Wang, X., and Tanaka Hall, T. M. (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein, *Nat. Struct. Biol. 8*, 141−145.
71. Storozhenko, S., Inze, D., Van Montagu, M., and Kushnir, S. (2001) *Arabidopsis* coactivator ALY-like proteins, DIP1 and DIP2, interact physically with the DNA-binding domain of the Zn-finger poly(ADP-ribose) polymerase, *J. Exp. Bot. 52*, 1375−1380.
72. Javerzat, J. P., Cranston, G., and Allshire, R. C. (1996) Fission yeast genes which disrupt mitotic chromosome segregation when overexpressed, *Nucleic Acids Res. 24*, 4676−4683.
73. Liker, E., Fernandez, E., Izaurralde, E., and Conti, E. (2000) The structure of the mRNA export factor TAP reveals a *cis* arrangement of a non-canonical RNP domain and an LRR domain, *EMBO J. 19*, 5587−5598.
74. Gruter, P., Tabernero, C., von Kobbe, C., Schmitt, C., Saavedra, C., Bachi, A., Wilm, M., Felber, B. K., and Izaurralde, E. (1998) TAP, the human homologue of Mex67p, mediates CTE-dependent RNA export from the nucleus, *Mol. Cell 1*, 649−659.
75. Rodrigues, J. P., Rode, M., Gatfield, D., Blencowe, B. J., Carmo-Fonseca, M., and Izaurralde, E. (2001) REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus, *Proc. Natl. Acad. Sci. U.S.A 98*, 1030−1035.
76. Wichmann, I., Garcia-Lozano, J. R., Respaldiza, N., Gonzalez-Escribano, M. F., and Nunez-Roldan, A. (1999) Autoantibodies to transcriptional regulation proteins DEK and ALY in a patient with systemic lupus erythematosus, *Hum. Immunol. 60*, 57−62.
77. Portman, D. S., O'Connor, J. P., and Dreyfuss, G. (1997) YRA1, an essential *Saccharomyces cerevisiae* gene, encodes a novel nuclear protein with RNA annealing activity, *RNA 3*, 527−537.
78. Kataoka, N., Yong, J., Kim, V. N., Velazquez, F., Perkinson, R. A., Wang, F., and Dreyfuss, G. (2000) Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm, *Mol. Cell 6*, 673−682.
79. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.

BI034062O