Characterization of a Family of RanBP2-Type Zinc Fingers that Can Recognize Single-Stranded RNA

Cuong D. Nguyen†, Robyn E. Mansfield†, Wilfred Leung†, Paula M. Vaz, Fionna E. Loughlin, Richard P. Grant and Joel P. Mackay*

School of Molecular Bioscience, University of Sydney, Sydney NSW 2006, Australia

Received 24 September 2010; received in revised form 24 December 2010; accepted 28 December 2010
Available online 20 January 2011

Edited by M. F. Summers

Keywords:
ZRANB2; EWS; TLS/FUS; RBM5; TEX13A

The recognition of single-stranded RNA (ssRNA) is an important aspect of gene regulation, and a number of different classes of protein domains that recognize ssRNA in a sequence-specific manner have been identified. Recently, we demonstrated that the RanBP2-type zinc finger (ZnF) domains from the human splicing factor ZnF Ran binding domain-containing protein 2 (ZRANB2) can bind to a sequence containing the consensus AGGUAA. Six other human proteins, namely, Ewing’s sarcoma (EWS), translocated in liposarcoma (TLS)/FUS, RNA-binding protein 56 (RBP56), RNA-binding motif 5 (RBM5), RNA-binding motif 10 (RBM10) and testis-expressed sequence 13A (TEX13A), each contains a single ZnF with homology to the ZRANB2 ZnFs, and several of these proteins have been implicated in the regulation of mRNA processing. Here, we show that all of these ZnFs are able to bind with micromolar affinities to ssRNA containing a GGU motif. NMR titration data reveal that binding is mediated by the corresponding surfaces on each ZnF, and we also show that sequence selectivity is largely limited to the GGU core motif and that substitution of the three flanking adenines that were selected in our original selection experiment has a minimal effect on binding affinity. These data establish a subset of RanBP2-type ZnFs as a new family of ssRNA-binding motifs.

© 2011 Elsevier Ltd. All rights reserved.

Introduction

The processes of transcription, transport, processing, storage, translation and degradation of messenger RNAs are all regulated by specific proteins, many of which contain discrete domains for the recognition of single-stranded RNA (ssRNA). Although less is understood about the recognition of ssRNA in comparison to double-stranded DNA, the recent determination of a number of structures has begun to provide insight into several classes of RNA-binding domain (see Ref. 1 for a review). RNA recognition motifs (RRMs; see Ref. 2 for a review), K-homology domains, Pumilio repeat domains and Tis11d-type zinc fingers (ZnFs) recognize ssRNA through a variety of mechanisms. For example, RRMs display a substantial amount of diversity in their interactions with RNA; different family
members recognize between four and eight bases using a combination of side-chain- and backbone-mediated hydrogen bonds, as well as stacking interactions involving aromatic side chains. In contrast, Tis11d binds to AU-rich sequences using almost exclusively interactions mediated by the protein backbone.

ZnF Ran binding domain-containing protein 2 (ZRANB2) (ZNF265, Zis) is a splicing factor that contains an arginine-serine-rich domain and two RanBP2-type ZnFs. It has been shown to interact with other components of the splicing apparatus, namely, U2AF65 and U1-70K, and to alter the splicing pattern of a reporter gene (Tra2β1) in splicing assays.6 Recently, we used Systematic Evolution of Ligands by EXponential enrichment (SELEX) to demonstrate that each of the ZnFs is capable of recognizing ssRNA that carries the consensus sequence AGGUAA.7 We also determined the three-dimensional structure of a complex formed by the second ZnF of ZRANB2 (ZRANB2-F2) and an oligonucleotide carrying the consensus sequence. This structure (Fig. 1a) revealed that the molecular basis for RNA recognition was distinct from that of previously described RNA-binding domains.

In the structure, two sequential guanine bases are recognized by “double-headed” hydrogen bonds from arginine side chains (R81 and R82), and these bases also stack on either side of a tryptophan side chain (W79), forming a unique Gua-Trp-Gua ladder. The uridine forms hydrogen bonds with two asparagine side chains (N76 and N86). It was notable that no base-specific interactions were observed to any of the three adenines that flank the GGU sequence, despite these bases clearly being selected in the SELEX experiment.

RanBP2-type ZnFs are found in organisms as diverse as fungi, plants and mammals, and analysis of the human genome reveals ∼30 other proteins that contain these domains. In several cases, these domains have been identified as protein-binding modules. For example, RanBP2 ZnFs from the nuclear pore complex proteins RanBP2 (Nup358) and Nup153 have been shown to bind RanGDP,8,9 while RanBP2 ZnF from the nuclear protein localization protein 4 homologue (NPL4) facilitates binding to ubiquitin.10,11 Alignment of the sequences of RanBP2 ZnFs reveals that a subset of the residues that are important for RNA recognition in ZRANB2-F2 is also present in at least six other human RanBP2 ZnFs (Fig. 1b), suggesting that these RanBP2 ZnFs might bind RNA at related sites. These proteins are translocated in liposarcoma (TLS or FUS),12 Ewing’s sarcoma (EWS),13 RNA-binding motif 5 (RBM5 or LUCA-15),14 RNA-binding motif 10 (RBM10),15 RNA-binding protein 56 (RBP56, TAF15 or TAF168),16 and testis-expressed sequence 13A (TEX13A).17

Most of these proteins have been implicated in transcriptional regulation and/or RNA processing. RBM5 is a putative tumor suppressor that associates with pre-spliceosomal complexes18 and has recently been shown to regulate splice site pairing in apoptosis-related genes such as Fas.19 RBM10 shares ∼50% sequence identity with RBM5 and is therefore likely to have a similar function. TLS/FUS was discovered by virtue of its N-terminal region forming part of a fusion oncoprotein associated with liposarcomas,12 and it has been purified as part of a heterogeneous nuclear ribonucleoprotein complex that is involved in both splicing and RNA transport.20 Most recently, mutations in TLS have been shown to be associated with the motor neuron disease amyotrophic lateral sclerosis.21 EWS and RBP56, which share a similar domain structure with TLS (Fig. 1c) and together form the TET family of proteins, both co-purify with the TFIID pre-initiation complex and both contribute strong transactivation domains to oncogenic fusion proteins.13,22

With the exception of TEX13A, these proteins all contain RNA-recognition motifs (RRMs) and have been shown to bind to ssRNA sequences that are rich in Gua and/or Ura.23-27 In general, however, the domain(s) responsible for the binding activity have not been delineated. In the case of TLS, SELEX data revealed a preference for GGGU sequences26 and 15N-heteronuclear single quantum coherence (HSQC) titrations showed that the ZnF, but not the RRM, was responsible for this binding activity.28 Despite a growing body of data on each of these proteins and evidence of their involvement in a variety of gene regulatory functions, their mechanisms of action are generally not well defined at the biochemical level. Here, we demonstrate that the RanBP2-type ZnFs of all of these proteins bind in a stoichiometric and sequence-specific manner to ssRNA containing the sequence AGGUAA, forming complexes with well-defined structures, and that their affinities for this sequence are comparable to those observed for ZRANB2:RNA interactions. Chemical shift perturbation data further reveal that the same surface is used by each domain to contact RNA. Our data demonstrate that these domains together constitute a distinct class of ssRNA-binding domain and should assist in defining the functions of the associated proteins.

Results

Assessment of the sequence specificity of ZRANB2

Our SELEX data for ZRANB2 clearly showed a preference for adenines in the positions flanking the core GGU sequence. However, fluorescence
anisotropy binding data for both the first ZnF of ZRANB2 (ZRANB2-F1 and ZRANB2-F2) indicated that changing each of these adenines to cytosine had a minimal effect on the affinity of the ZnF for RNA. In order to obtain a more comprehensive view of the binding specificity of ZRANB2, we assessed the binding of ZRANB2-F2 to a series of oligonucleotides in which we altered each of the first five bases in the AGGUAA sequence to each of the other three possible bases. The data shown in Fig. 2 indicate that no single-base mutants bind to ZRANB2-F2 more tightly than the AGGUAA consensus sequence. These data also confirm that a significant preference exists for guanine at positions 2 and 3 and for uridine or guanine at position 4. In contrast, there is very little sequence specificity for the three flanking adenine bases that were selected in our SELEX experiment.
The RanBP2-type ZnFs of several human proteins can bind ssRNA

We next expressed the ZnF domains from EWS, TLS, RBP56, RBM5, RBM10 and TEX13A as glutathione S-transferase (GST)-fusion proteins and purified the fusion proteins using glutathione-affinity chromatography. A portion of each protein was cleaved from the GST, and a one-dimensional 1H NMR spectrum was recorded in order to assess the folding of each domain. These spectra (not shown) confirmed that each domain took up a well-ordered conformation in solution.

We then used fluorescence anisotropy titrations to test the ability of each domain to bind to a 17-nt ssRNA containing a single AGGUAA motif. The ubiquitin-binding RanBP2 ZnF from NPL4, which lacks the conserved RNA-binding residues (Fig. 1b), was examined in parallel. Figure 3 shows that, with the exception of NPL4, all of the domains bound to RNA with affinities that were similar to those of the individual ZRANB2 ZnFs (0.4–3.0×10^6 M^-1). In contrast, no binding was observed when a 17-nt oligonucleotide comprising only cytosines was used, indicating that the interaction is sequence specific. Furthermore, GST alone did not bind to either of the RNA sequences with an appreciable affinity.

Each domain forms a specific, well-ordered 1:1 complex with RNA

We next used NMR spectroscopy to examine these ZnF:RNA interactions. We prepared uniformly 15N-labeled versions of EWS, RBP56, RBM5, RBM10 and TEX13A, cleaving each domain from GST and purifying it by gel-filtration chromatography. Each domain yielded a high-quality 15N-HSQC spectrum. We then carried out chemical shift perturbation experiments, recording 15N-HSQC spectra, as either a 6-nt (AGGUAA) or a 10-nt (CCAGGUAAAG) ssRNA was titrated into the protein. Figure 4 shows that each titration induced substantial chemical shift changes and that these changes were restricted to a subset of residues on each ZnF. The spectra obtained following the addition of 1–1.25 molar equivalents of RNA were in general sharp and well dispersed, consistent with the formation of well-ordered monomeric complexes.

In each case, the titration indicated that formation of the complex was in intermediate to fast exchange on the chemical shift timescale, consistent with the measured binding affinities. Thus, some signals simply shifted in position over the course of the titration, whereas others disappeared during the titration. In each case, approximately the same number of signals that had disappeared was observed to reappear following the addition of 1 molar equivalent of RNA, and no significant changes were observed following the further addition of 1.25 molar equivalents of RNA (full sets of titration spectra for each ZnF are provided in Fig. S1–5).

The RNA interaction surface is conserved between each RanBP2 ZnF

In order to compare the mode of binding of each of these ZnFs with that observed previously for ZRANB2, we prepared uniformly 15N,13C-labeled ZnFs and recorded backbone triple-resonance experiments to make assignments of backbone nuclei (Fig. 4). Assignments of HN, N, Cα and Cβ resonances were made for all but 1–3 residues for each ZnF alone (a small number of residues showed no assignable signal in the corresponding 15N-HSQC most likely due to local dynamic processes) and for 32 of the expected 36 residues for the EWS ZnF:RNA complex (6 residues in intermediate exchange did not reappear following the addition of 1 molar equivalent of RNA). Assignments of HN and N resonances were also made for the majority of residues in the RNA-bound state for the ZnFs of RBP56, RBM5, RBM10 and TEX13A by following the migration of fast-exchange signals throughout the RNA titrations. From these assignments, we calculated weighted average chemical shift changes (for HN and N) for each residue (on the left of Fig. 5). On the right of Fig. 5, changes that are greater than 1 SD above the mean change are mapped onto the homology models of the EWS, RBP56, RBM5, RBM10 and TEX13A ZnFs generated using SWISS-MODEL and the structure of ZRANB2-F2. Template structures for the homology models were
chosen automatically by SWISS-MODEL and corresponds to ZRANB2-F1 [Protein Data Bank (PDB) ID: 1N0Z] in the case of RBM5, RBM10 and TEX13A and to ZRANB2-F2 in complex with RNA (PDB ID: 3G9Y) in the case of EWS and RBP56. Overall, the RNA-binding surface defined by these data resembles that observed for ZnFs from ZRANB2 (Fig. 5f).

It is notable, however, that TEX13A ZnF displayed some significant chemical shift changes on the “back” side of the domain, perhaps indicating that this ZnF undergoes some conformational change upon binding to RNA (Fig. 5e).

**The RNA-binding specificity of TLS and RBM5**

In order to probe the specificity of the interaction for these proteins, we measured the effect of single-point mutations in the RNA sequence on the binding affinity of TLS ZnF. Figure 6a shows that changes in either of the two guanines (Gua → Cyt mutations) reduce the association constant by 3- to 4-fold, consistent with the data for ZRANB2-F1 and ZRANB2-F2. Given the preference for GGUG sequences demonstrated for TLS in published SELEX data, we also tested the binding of this domain to a sequence containing an AGGUGA motif. As shown in Fig. 6a, the affinity for this interaction was approximately twofold lower than that of the interaction with AGGUAA. Mutations of the flanking adenines did in two cases also result in approximately twofold reductions in affinity (A1C and A5G mutations), suggesting that there might be some protein contacts made by these nucleotides.
Characterization of RanBP2-Type Zinc Fingers

Fig. 4 (legend on next page)
Comparison of the amino acid sequences given in Fig. 1b reveals that, in a subset of RNA-binding RanBP2-type ZnFs (RBM5, RBM10 and TEX13A), the two asparagines that contact Ura4 (N76 and N86 in ZRANB2-F2, Fig. 1a) are changed to hydrophobic residues. To assess whether these changes have a significant impact on sequence specificity, we measured the binding of RBM5 to oligonucleotides containing each of the four bases at position 4 (AGGNA). The data in Fig. 6b (and comparison with Fig. 2) show that these amino acid substitutions significantly affect the sequence specificity of RBM5, which displays a 1.5-fold preference for guanine over adenine and cytosine are 2- and 8-fold, respectively.

Comparison of the RBM5 amino acid sequence with that of ZRANB2 also reveals that the arginine that contacts Gua2 in the latter protein (R81 in ZRANB2-F2) is substituted with lysine in RBM5. Surprisingly, measurement of the affinity of RBM5 for RNA sequences containing substitutions at position 2 (ANGUAA) shows that the sequence specificity imparted by lysine is similar to that achieved by arginine in this position (Fig. 6b and comparison with Fig. 2).

Discussion

A family of RNA-binding RanBP2-type ZnFs

We show here that the RanBP2-type ZnFs from the human proteins TLS, EWS, RBP56, RBM5, RBM10 and TEX13A are bona fide ssRNA-binding proteins and that they can bind with micromolar affinity to sequences containing a central GGU motif. These findings corroborate published biochemical data that point toward a nucleic acid binding capability for most of this set of proteins. Thus, full-length RBP56, RBM5 and RBM10 were shown to bind in vitro to ssRNA containing a poly-G stretch, whereas SELEX and NMR data revealed that TLS can recognize GGUG sequences.

The TET family proteins TLS, EWS and RBP56 share similar domain structures, and it has previously been suggested that RanBP2 ZnFs in these proteins might function as DNA-binding domains. We previously showed that ZRANB2 ZnFs do not interact appreciably with double-stranded DNA containing the AGGUAA consensus motif. Our data further showed that although these domains can interact with single-stranded DNA, the interaction is likely to be substantially weaker than the interaction with ssRNA. Based on these results and our current data, we conclude that it is more likely that the RanBP2 domains from each of the proteins that we have examined function to recognize ssRNA.

It is interesting to note that S439 of TLS can be phosphorylated. This residue sits in the center of the RNA-binding surface of this domain, suggesting that phosphorylation might act as a mechanism by which RNA binding is regulated in this protein.

Conservation of RNA-binding RanBP2-type ZnFs

BLAST searches carried out using each of the ZnFs in this study revealed that the domains are present in the apparent paralogues of these proteins in a number of other organisms (Fig. S6). The RNA-binding residues of TLS, EWS and RBP56 are conserved in mammals, birds, amphibians, bony fish and insects, indicating an ancient and highly conserved RNA-binding activity. Highly related domains are also observed in organisms as diverse as the parasite Schistosoma japonicum and the nematode Caenorhabditis elegans, as well as in the plants Arabidopsis thaliana and Oryza sativa. It is also notable that all of these related proteins additionally contain an RRM, one of the best-characterized ssRNA-binding domains, suggesting that these two domain classes can combine to fine-tune target specificity.

Close examination of the sequences of the domains in these other organisms reveals several changes in the putative RNA recognition residues—some conservative and some relatively nonconservative—and it will be of interest to establish whether these changes subtly alter RNA sequence specificity.

The design of novel RNA-binding proteins

The recent development of ZnF proteins with tailored DNA-binding specificity has resulted in a wide range of research and clinical applications (see Refs. 32–34 for reviews). Tandem repeats of three or more of these ZnF domains have been fused to a variety of effector domains to create new proteins that can act, for example, as transcriptional regulators or site-specific nucleases. The success of these proteins has been built on the properties of the classical ZnF, including small size, stability to...
Fig. 5 (legend on next page)
mutagenesis and the fact that interactions with the DNA target are mediated exclusively by side-chain contacts, allowing a range of binding specificities to be generated on the same scaffold. RanBP2-type ZnFs exhibit similar properties and as such display potential as a scaffold for the design of sequence-specific RNA-binding proteins that could be used to target ssRNA in the cell. The addition of effector domains to an RNA-binding domain with tailored specificity could yield designer translation factors, splicing factors or probes of RNA localization, and it will be of great interest to see whether an application of this type can be realized in the future.

Materials and Methods

Cloning, expression and purification of RanBP2 ZnF constructs

The RanBP2 ZnF domain from the human proteins TLS (residues 422–453; NM_004960), EWS (514–550; NM_013986), RBP56 (349–388; NM_005778), RBM5 (176–213; NM_005676), TEX13A (370–409; NM_031274) and NPL4 (580–608; NM_017921) was amplified from a K562 cDNA library and inserted into the plasmid pGEX-2T (GE Healthcare). Each construct was expressed in Escherichia coli BL21 cells, as a fusion with GST and cell pellets were lysed by sonication in buffer containing 1% (v/v) Triton X-100, 3 mM phenylmethanesulfonyl fluoride and phosphate-buffered saline, pH 7.4. 15N- and 15N,13C-labeled RanBP2 ZnFs were prepared following the protocol of Cai et al.35 Proteins used for fluorescence anisotropy experiments were purified by glutathione-affinity chromatography as a fusion with GST, while proteins used for NMR experiments were cleaved from GST using thrombin and further purified by size-exclusion chromatography (HiLoad 16/60 Superdex 30).

RNA

RNA oligonucleotides were purchased from Dharmacon and were deprotected prior to use according to the instructions provided by the manufacturer. For fluorescence anisotropy experiments, 17-nt oligonucleotides that were based on the sequence GCAACCAGGUAAAGUCU were used. For NMR measurements, 6- or 10-nt sequences were used (AGGUAA or CCAGGUAAAG).

Fluorescence anisotropy titrations

GST-tagged RanBP2 ZnFs and RNA were dialyzed into buffer containing 10 mM Tris (pH 8), 50 mM KCl, 5 mM
MgCl₂ and 1 mM DTT made with diethyl-pyrocarbonate-treated water. All buffers and equipment were kept RNase free at all times, and RNasin (Promega) was added to the protein solution to inhibit RNase activity. Fluorescence anisotropy titrations were performed at 25 °C, with a slit width of 10 nm and with excitation and detection at 495 and 520 nm, respectively; data were averaged over 15 s. In each titration, the fluorescence anisotropy of a solution of 50 nM fluorescein-tagged RNA containing 0.05 mg/mL heparin (to reduce nonspecific binding) was measured as a function of added protein concentration. Binding data were fitted to a simple 1:1 binding model by nonlinear least-squares regression. Each titration was performed three times, and the final affinity was taken as the mean of these measurements.

NMR spectroscopy

For chemical shift perturbation experiments, purified 15N-labeled RanBP2 ZnFs were concentrated to 70–250 μM. The 15N,13C-labeled RanBP ZnFs were concentrated to 0.3–1.5 mM. All NMR samples contained 5% D₂O and 20 μM 2,2-dimethyl-2-silapentane-5-sulfonic acid as a chemical shift reference. All experiments were run at 298 K on a 600- or 800-MHz Bruker Avance spectrometer equipped with a cryoprobe. For titrations, proteins and 10-nt RNA were dialyzed into 25 mM sodium phosphate, pH 6.5, containing 100 mM NaCl and 1 mM DTT. The 6-nt RNA was buffer exchanged into water using a mini Quick Spin Oligo Column (Roche), lyophilized and resuspended in the same buffer as the protein. 13N-HSQC spectra were recorded for the ZnF alone and following the addition of typically 0.25, 0.5, 1 and 1.25 molar equivalents of RNA (exact titration points for each ZnF are specified in Figs. S1–S5). In order to assign backbone resonances of the RanBP2 ZnFs, we recorded HNCACB and CBCA(CO)NH experiments. The EWS ZnF, these spectra were also recorded in the presence of 6-nt RNA.

Accession numbers

Chemical shift assignments for HN, N, Cα and Cβ nuclei of EWS, RBP56, RBM5, RBM10 and TEX13A ZnFs have been deposited in the BioMagResBank under accession IDs 17133 (EWS), 17380 (RBP56), 17387 (RBM5), 17386 (RBM10) and 17385 (TEX13A).

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2010.12.041

Acknowledgements

We thank Ann Kwan and Bill Bubb for expert maintenance of the NMR facility. R.E.M., P.M.V. and F.E.L. were supported by Australian Postgraduate Awards. This work was supported in part by a Project Grant from the National Health and Medical Research Council of Australia (to J.P.M.).

References

Characterization of RanBP2-Type Zinc Fingers


