First publ. in: Journal of Molecular Biology ; 351 (2005), 4. - S. 718-730
DOI : 10.1016/j.jmb.2005.06.032

Solution Structure of the N-terminal Zinc Fingers of the *Xenopus laevis* double-stranded RNA-binding Protein ZFa

Heiko M. Möller, Maria A. Martinez-Yamout, H. Jane Dyson and Peter E. Wright*

Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla CA 92037, USA

Several zinc finger proteins have been discovered recently that bind specifically to double-stranded RNA. These include the mammalian JAZ and wig proteins, and the seven-zinc finger protein ZFa from *Xenopus laevis*. We have determined the solution structure of a 127 residue fragment of ZFa, which consists of two zinc finger domains connected by a linker that remains unstructured in the free protein in solution. The first zinc finger consists of a three-stranded β-sheet and three helices, while the second finger contains only a two-stranded sheet and two helices. The common structures of the core regions of the two fingers are superimposable. Each finger has a highly electropositive surface that maps to a helix-kink-helix motif. There is no evidence for interactions between the two fingers, consistent with the length (24 residues) and unstructured nature of the intervening linker. Comparison with a number of other proteins shows similarities in the topology and arrangement of secondary structure elements with canonical DNA-binding zinc fingers, with protein interaction motifs such as FOG zinc fingers, and with other DNA-binding and RNA-binding proteins that do not contain zinc. However, in none of these cases does the alignment of these structures with the ZFa zinc fingers produce a consistent picture of a plausible RNA-binding interface. We conclude that the ZFa zinc fingers represent a new motif for the binding of double-stranded RNA.

*Corresponding author

Keywords: NMR; helix-loop-helix; helix-turn-helix; JAZ; wig-1

Introduction

In recent years double-stranded RNA (dsRNA) has attracted great interest due to its role in defense against viruses, processing of cellular RNA and post-transcriptional gene regulation known also as RNA interference and gene silencing.[1–4] Unlike the *B* and *Z* forms typical of double-stranded DNA (dsDNA), dsRNA is found commonly in the *A* form, characterized by a deep, narrow major groove and a broad, shallow minor groove.[5] The narrow major groove impedes contact between dsRNA-binding proteins and nucleobases so that sequence specificity such as that observed for DNA-binding proteins is difficult to implement. Indeed, no report of sequence-specific recognition of dsRNA has been published.[6] Recognition of dsRNA will most likely prove to be mainly shape-dependent, with structural irregularities such as bulges, loops, and mismatches being of major importance.[6–8]

The wide variety of functions being described for dsRNA would seem to implicate a variety of structures for interacting proteins,[9] but only a few dsRNA-binding motifs have been characterized, primarily the dsRNA-binding motif (dsRBM), a well-characterized domain consisting of about 70 amino acid residues with an αββββα fold.[10,11] Recently, a series of zinc finger proteins have been described that appear to interact specifically with dsRNA. Zinc finger proteins are well

Present address: Heiko M. Möller, Fachbereich Chemie, Universität Konstanz, 78457 Konstanz, Germany.

Abbreviations used: dsRNA, double-stranded RNA; dsDNA, double-stranded DNA; ssRNA, single-stranded RNA; HSQC, heteronuclear single quantum coherence; NOESY, nuclear Overhauser effect spectroscopy; COSY, correlated spectroscopy.

E-mail address of the corresponding author: wright@scripps.edu

known to interact with dsDNA and single-stranded RNA (ssRNA). Here, we describe the structure of the first two zinc fingers of the zinc finger domain of the *Xenopus laevis* dsRNA-binding protein (dsRBP-ZFa),[12] a protein that typifies a new class of zinc fingers that show high affinity for dsRNA, while binding much more weakly to dsDNA or single-stranded nucleic acids (RNA or DNA).

dsRBP-ZFa consists of an acidic C terminus and a basic N-terminal domain containing seven zinc finger motifs of the C_2H_2 type. In gel mobility-shift assays, dsRBP-ZFa showed binding to dsRNA and RNA/DNA hybrids. Nanomolar affinity of the 456 residue zinc finger domain was observed to a dsRNA epitope calculated to be, at most, 25 bp in length.[12] A construct consisting of the three N-terminal zinc fingers also binds dsRNA, but with an estimated binding epitope of only 12 bp.[13] The amino acid sequence of dsRBP-ZFa is atypical, in that the individual zinc finger domains are connected by very long linkers (34–44 amino acid residues). The majority of dsDNA-binding zinc finger proteins have linkers of six to eight amino acid residues. In addition, the TGEKP-motif that is highly conserved in other C_2H_2 zinc fingers and that stabilizes the domains in the bioactive conformation upon binding to DNA,[14] is absent from the linker sequences of dsRBP-ZFa. Another difference with respect to the classical C_2H_2 zinc fingers is the increased spacing between the zinc-coordinating histidine residues (five instead of the usual three or four residues).

While dsRBP-ZFa is mainly interesting as the prototypical member of a new class of zinc finger proteins displaying unique specificity towards dsRNA, its function in *X. laevis* remains unclear.[12] Homologs of dsRBP-ZFa, designated wig-1[15] and PAG608,[16] found in humans and rodents, are target genes of the tumor suppressor p53 and, together with the mammalian homolog JAZ, are involved in the apoptotic response.[15–17] All of these proteins bind specifically to dsRNA with their zinc finger domains, and essentially require this capability in order to function within the apoptotic network. A sequence alignment of the zinc finger domains of ZFa, JAZ and wig proteins is shown in Figure 1. The sequence similarity between these three classes of proteins is greatest in the first two and the fourth zinc finger domains. The domain of wig1 that corresponds to the third zinc finger of ZFa and JAZ appears to lack the zinc ligands, although some of the hydrophobic residues in this region appear to retain homology with the sequences of ZFa and JAZ. Interestingly, only ZFa contains the fifth, sixth and seventh zinc fingers; these appear to have arisen through multiple duplication of the fourth zinc finger.[12] If this is so, then it is possible that ZFa represents the equivalent in the frog of the mammalian JAZ protein.

No three-dimensional structure of the dsRNA binding zinc finger proteins has been reported. We present here the solution structure of the first two zinc fingers of dsRBP-ZFa. The two domains each show a zinc-dependent αβ fold, somewhat similar to each other and to canonical zinc finger structures. Comparison of the ZFa zinc finger structures with published structures of primarily DNA-binding zinc fingers reveals intriguing differences in sequence and structure that may account for the major differences in binding specificity between these protein types.

Results

Design and preparation of ZFa zinc finger constructs

Inspection of the sequence alignment (Figure 1) reveals excellent homology between ZFa, JAZ and wig proteins in the first two zinc finger domains, though not in the linkers between the fingers. In order to provide insight into the mechanism of dsRNA recognition by this new class of zinc finger proteins, we prepared three constructs corresponding to residues 1–74, 1–128 and 1–205 of dsRNAbp ZFa. Expression of constructs 1–128 and 1–205 with the wild-type sequence resulted in each case in the production of two well-defined species in a ratio of about 1:1, a full-length product and a smaller product. Analysis by matrix-assisted laser desorption/ionization mass spectrometry for the 1–128 construct showed that the mass of the full-length product was consistent with residues 2–128. The mass of the truncated product was not consistent with any likely proteolytic cleavage of the longer constructs, but was consistent with residues 2–79 plus a valine residue. Inspection of the DNA sequence revealed that this product was possible if an extra base was inserted into a stretch of seven adenine residues within the DNA sequence of the construct. Two silent mutations, at codons 77 and 78 (GCA AAA → GCG AAG) in the ZFa DNA sequence were designed to interrupt the long sequence of adenine residues. Expression of the mutant construct resulted in elimination of the by-product and the production of 100% of the full-length product.

Since the ^{15}N-heteronuclear single quantum coherence (HSQC) spectra of the three ZFa constructs (data not shown) were superimposable over the common residues, indicating no interaction between the zinc fingers, and the minimal high-affinity dsRNA-binding site of the homologous protein JAZ has been mapped to two sequential fingers,[17] ZFa 1-128 (with the expressed protein containing residues 2–128) was chosen for structure determination in solution by NMR. Analysis of the amino acid sequence of this portion of the protein already suggests two C_2H_2 zinc coordination sites, and indeed, folded dsRBP-ZFa2-128 was obtained only in the presence of zinc. A comparison of the matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry of the pure folded protein gave a mass difference of

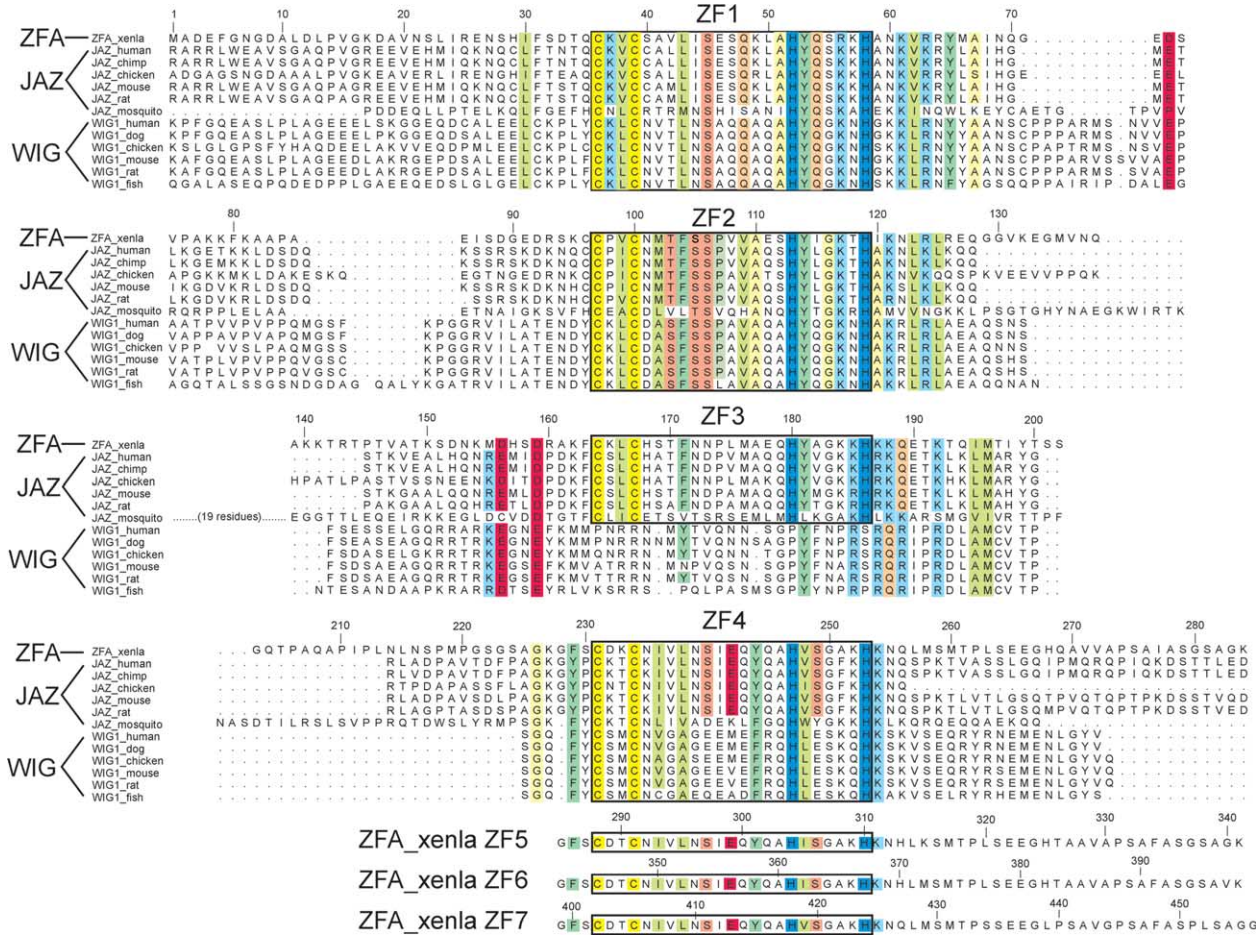

Figure 1. Amino acid sequence alignment of the ZFa protein from *Xenopus laevis* with the JAZ and wig proteins from a number of species. The sequences are arranged to emphasize the homology and identity between fingers. Ligand cysteine and histidine residues are outlined in yellow and blue, respectively, conserved hydrophobic, aromatic, positive, negative, hydrophilic and small residues are outlined in green, blue-green, light blue, red, pink and light yellow, respectively. Numbers correspond to the ZFa sequence. The JAZ and wig proteins have extensions at the N terminus that are not shown.

128 kDa, showing the presence of two zinc ions. The circular dichroism spectrum of ZFa (Figure 2) is typical for a folded protein containing helix (minima at 208 nm and 222 nm). Addition of EDTA to the solution results in loss of these signals, an indication that the zinc ions are required for folding of the protein.

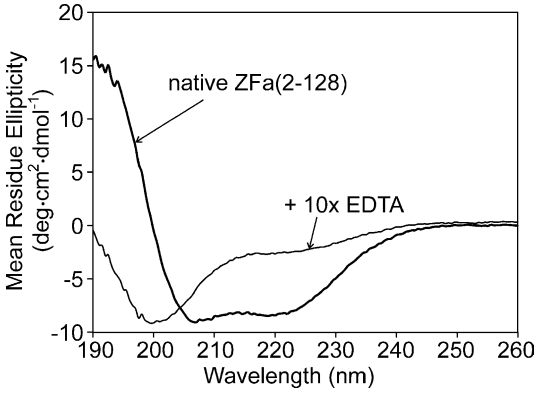

Figure 2. UV CD spectra of dsRBP-ZFa2-128 (thick line) and after removal of zinc with a tenfold molar excess of EDTA (thinner line).

Resonance assignment

Nearly complete resonance assignments for dsRBP-ZFa2-128 (97% backbone, 82% side-chain) were obtained by analysis of multidimensional heteronuclear NMR spectra (see Experimental Procedures). The ^{15}N HSQC spectrum of dsRBP-ZFa2-128 is shown in Figure 3. Due to the high content of arginine and lysine residues, some side-chain resonances, especially those located in the linker between the zinc fingers (residues 69–92) could not be assigned uniquely. Significant crowding of amide resonances is observed, due to the presence of α-helix and the lengthy unstructured regions at the N terminus (11 residues) and in the linker region (24 residues). In addition, the scarcity of aromatic residues, including the complete absence of tryptophan, results in limited dispersion of the side-chain resonances. For example, 29/58

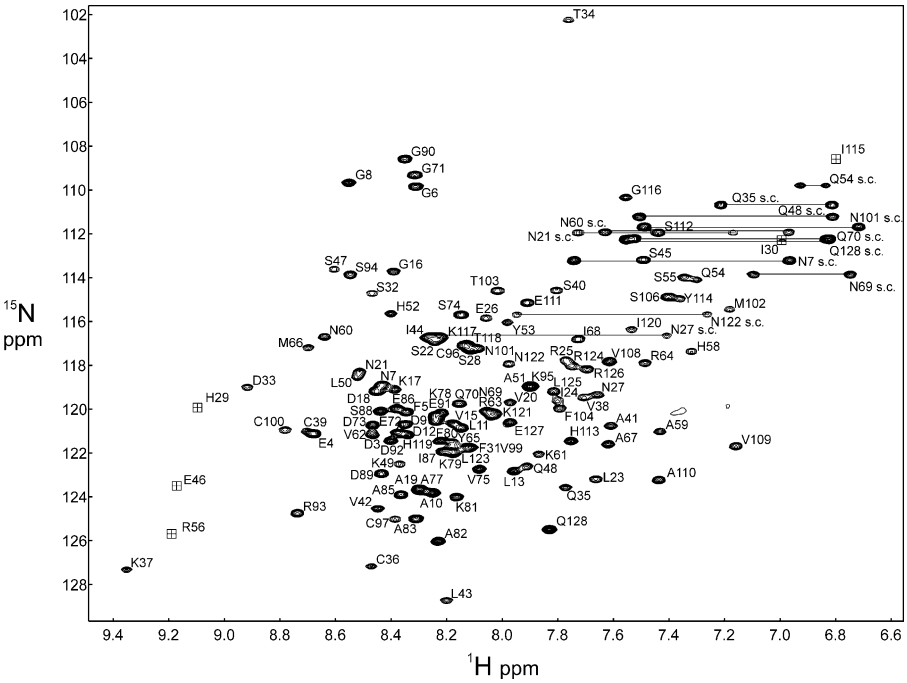

Figure 3. ^{1}H–^{15}N HSQC spectrum of ZFa2-128. Backbone amide crosspeaks are labeled, and side-chain amide groups are indicated with a horizontal bar. Crosspeaks marked with boxes are weak, probably due to solvent exchange broadening.

(50%) of methyl resonances are observed between 0.8 ppm and 1.0 ppm.

Structure determination

NOE distance constraints were derived from ^{15}N and ^{13}C-separated nuclear Overhauser effect spectroscopy (NOESY)-HSQC spectra using the CANDID module of the CYANA software package.[18,19] In addition, torsion angles calculated from ^{3}J-coupling constants were used as input for structure calculations. Backbone ϕ torsion angle restraints were derived from a 3D HNHA spectrum.[20] In regions with confirmed regular secondary structure (from the pattern of medium-range NOEs, $^{3}J_{HNHA}$ coupling constants and secondary chemical shifts of C^{α}, H^{α} and CO), the ψ torsion angles predicted by TALOS were used as additional constraints.[21]

Stereospecific assignments were made using coupling constants derived from HNHA and HACAHB-correlated spectroscopy (COSY) spectra and the local NOE pattern. For isoleucine, threonine and valine $^{3}J_{C\gamma N}$ and $^{3}J_{C\gamma CO}$ coupling constants, measured by spin-echo difference ^{13}C-CT-HSQC spectra,[22,23] were used to obtain stereospecific assignments of methyl groups and χ_1 torsion angles.

Initial CANDID/CYANA calculations were carried out with NOESY peaklists and TALOS predictions as the only input. Criteria published for assessing the quality of CANDID calculations were fulfilled for both zinc-binding domains.[18,19] Once the global fold had been established, the full sets of NOE and dihedral angle restraints were applied to the CYANA/CANDID calculations and later to refinement with AMBER†. Throughout the calculations, the first zinc finger was better defined than the second, probably due to overlap of the resonances of finger II, together with lower-quality NOE data for the second domain caused by shorter correlation time and/or exchange phenomena. The final set of restraints was used for refinement with AMBER 8.[24] Final refinement was carried out with a generalized Born solvation model.[25,26] The statistics of the final structure calculation is summarized in Table 1.

Description of the 3D structure

The first two zinc finger domains of dsRBP-ZFa are structurally well defined from residue L13 to I68 (finger I) and from R93 to E127 (finger II). The N terminus from A2 to D12 and the linker between the two zinc-binding domains, from N69 to D92, are disordered. This is demonstrated by heteronuclear ^{1}H–^{15}N NOE measurements (Figure 4). The two zinc finger domains tumble completely independently, and there is no interdomain or domain-linker NOE contact.

Families of structures for the two zinc finger domains are shown in Figure 5(a) and a single structure of each showing elements of secondary structure in Figure 5(b). The first zinc finger adopts an $\alpha\beta\beta\beta\alpha\alpha$ fold. The N-terminal α-helix (residues K17–E26, helix α1) is connected through a type I turn to a three-stranded antiparallel β sheet

† http://amber.scripps.edu

Table 1. Structure statistics

A. *NMR constraints*	
Total unambiguous distance restraints	1142
Intraresidue (*i*, *i*)	383
Sequential (*i*, *i*+1)	224
Medium-range ($2 \leq \|i-j\| \leq 4$)	270
Long-range ($\|i-j\| > 4$)	265
Total dihedral angle restraints	216
ϕ	102
ψ	82
χ_1	32
B. *Ensemble statistics (20 structures)*	
Violation analysis	
Maximum distance violation (Å)	0.27
Maximum dihedral angle violation ordered region (deg.)	2
Maximum dihedral angle violation disordered region (deg.)	8
Energies	
Mean restraint violation energy (kcal mol^{-1})	5.7 ± 1.1
Mean AMBER energy (kcal mol^{-1})	-4751 ± 36
Mean deviation from ideal covalent geometry	
Bond lengths (Å)	0.01
Bond angles (deg.)	2.44 ± 0.07
rmsd from mean structure	
Backbone heavy atoms (N, C^{α}, C′, O)	0.39 (finger I) 0.58 (finger II)
All heavy-atoms	0.97 (finger I) 1.22 (finger II)
Ramachandran plot[52]	
Most-favorable regions (%)	87.4
Additionally allowed regions (%)	12.4
Generously allowed regions (%)	0.1
Disallowed regions (%)	0.1

rmsd values and Ramachandran statistics are given for structured regions (residues 13–68 (finger I) and 93–127 (finger II)) only.

(residues F31–V42) that leads into a helix-kink-helix motif (residues E46–Q54, helix α2 and R56–M66, helix α3). The first strand and the second strand of the β-sheet are connected by a type IV turn. A type I turn leads from the second strand to the third strand. The zinc ligands consist of C36 in the second strand, C39 in the turn, H52 at the end of helix α2 and H58 at the start of helix α3. Helix α1 packs against helix α2 with the helix axes perpendicular to each other. A portion of the N-terminal sequence, residues 13–16, while not part of helix α1, is relatively well defined, and makes NOE contacts with residues of helices α2.

The second zinc-binding domain of dsRBP-ZFa contains a subset of the secondary structure elements of finger I, with a two-stranded antiparallel β sheet (residues 95–104, analogous to strands 2 and 3 of finger 1), connected by a type 1 turn. The connection of the β-sheet to the helix-kink-helix motif (residues 105–115, helix α4; and 117–127, helix α5) occurs through a loop with a structure very similar to that of finger 1. Indeed, the backbone atoms of finger II superimpose on the corresponding part of finger I almost perfectly (backbone rmsd 0.969 Å between the two lowest-energy structures).

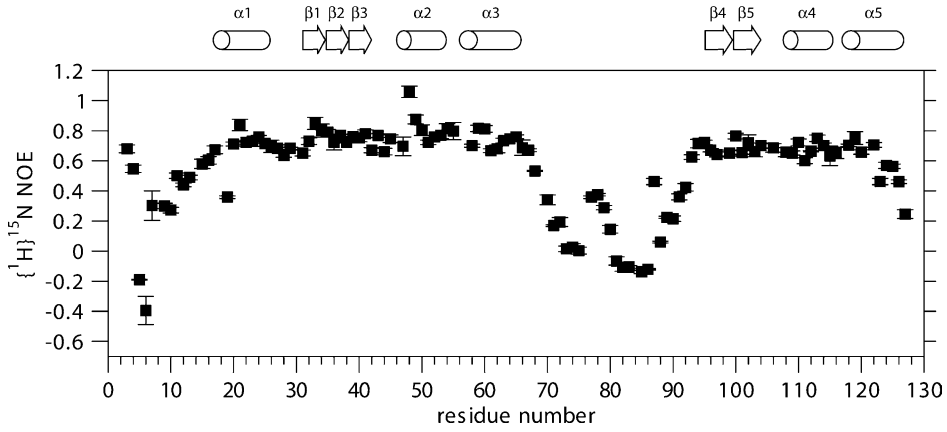

Figure 4. Heteronuclear ^1H–^{15}N NOEs measured for dsRBP-ZFa2-128. The cartoon on top of the Figure shows the location of secondary structure elements identified in the solution structure calculation.

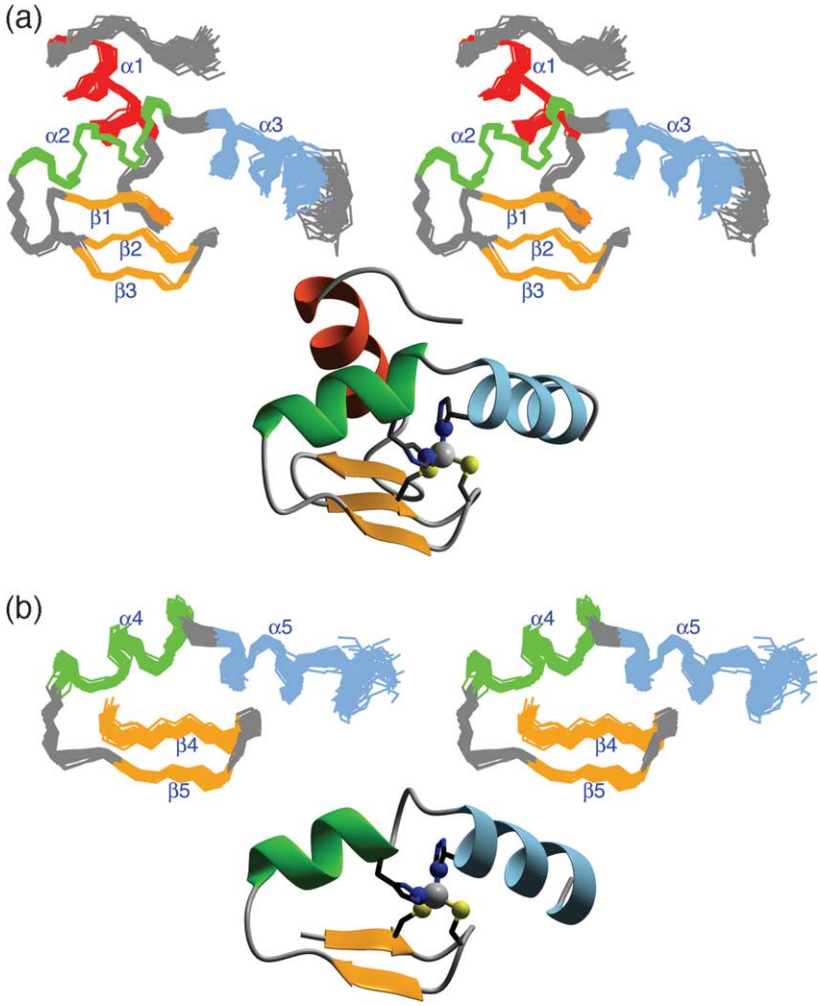

Figure 5. (a) Stereo view (upper panel) of a least-squares superposition of an ensemble of 50 structures of dsRBP-ZFa finger I, residues 13–68 and ribbon representation of the lowest-energy structure of finger I (lower panel). (b) Stereo view (upper panel) of a least-squares superposition of an ensemble of 50 structures of dsRBP-ZFa finger II, residues 95–127 and ribbon representation of the lowest-energy structure of finger II (lower panel). Loops are shown in gray, helix α1 (finger I) is colored red, homologous helices α2 (finger I) and α4 (finger II) are colored green, homologous helices α3 (finger I) and α5 (finger II) are colored blue, and the β-strands of both fingers are colored orange. The zinc coordination sphere is shown, with a gray sphere indicating the position of the zinc in each finger, and the coordinating sulfur atoms of C36 and C39 (finger I) and C97 and C100 (finger II) shown as yellow spheres. The coordinating nitrogen atoms, N^{ε} of H52 and H113 and N^{δ} of H58 and H119, are shown as blue spheres.

The high degree of similarity between the ββαα core of the first and second zinc finger domain of dsRBP-ZFa reflects hydrophobic packing of conserved residues in analogous positions in the two fingers. The hydrophobic core in the first zinc finger is formed by V38, L43, Y53 and V62. The side-chain of V38 is packed against K61, V62 and Y65 of helix α3, while L43 makes contact with Q48, K49, H52 and Y53 of helix α2. The analogous core in the second finger comprises V99, F104, Y114 and L123, with contacts between V99 and N122, L123 and (the aliphatic portion of) R126 (of helix α5), and between F104 and V109, A110, H113 and Y114 (of helix α4). In finger I, the side-chain of Y53 packs against both coordinating histidine residues, H52 and H58; the analogous residues in finger II are Y114, H113 and H119, respectively.

The β-sheets in both fingers are composed only of short strands with characteristic β dihedral angles, but are well defined according to a network of cross-strand NOEs, $^{3}J_{HNH\alpha}$ coupling constants and secondary chemical shifts. S40 and N101 perform analogous roles in the type I turn structure between the two β-strands. For each of these residues, the backbone populates a narrow region of the Ramachandran plot characterized by positive φ and ψ values around +60° and +30°, respectively, corresponding to the α_L minimum. This backbone conformation is supported by a very strong HN–H^{α} NOE, a large HN–H^{α} coupling constant and a negative secondary chemical shift of the β carbon atom, which are characteristic of positive backbone φ angles.[27,28] The backbone of the residue following a zinc-bound C-X-X-C motif frequently takes on an α_L configuration.[29]

The sequence alignment (Figure 1) suggests two C_2H_2 centers for zinc coordination, C36, C39, H52 and H58 in finger 1 and C97, C100, H113 and H119 in finger 2. However, this information was not used in initial structure calculations, in order to avoid bias associated with forcing a certain coordination scheme. From structure calculations based on NOEs and coupling constants alone, it became obvious that the side-chains of the putative coordinating residues are positioned suitably to allow zinc coordination. These calculations support zinc coordination *via* N^{δ} of H58 and H119, and *via* N^{ε} of H52 and H113. The ε,δ coordination was established unambiguously by acquisition of a 2D $^{15}N,^{1}H$-HMQC spectrum optimized for detecting two-bond and three-bond couplings between ring nitrogen atoms and non-exchangeable ring protons

of histidine residues.[30] The only possible ambiguity in the zinc coordination is the presence of an extra cysteine residue at position 96, which is next to the finger 2 zinc ligand C97. C96 is not well conserved (Figure 1), and invariably we observe the side-chain of C96 pointing into solution away from the zinc, and populating all three χ_1 rotamers in the structure calculations. We conclude that C96 is not involved in zinc coordination.

Comparison with other structures

The 3D structure searches were carried out with the DALI, SSM-SCOP, TOPS and VAST search engines.[31–34] Many zinc finger proteins were classified as similar to the structured domains of ZFa, because they contain the same ββα substructure. A particularly close resemblance was found between finger II and the zinc fingers of the friend-of-GATA (FOG) family and with canonical C_2H_2 zinc fingers such as those from TFIIIA and SWI5. There are familial resemblances to the structure of the human spliceosomal protein U1C. However, no zinc finger fold was found that contains all of the distinctive features of the first zinc-binding domain of ZFa; namely, the N-terminal α-helix, the third strand of the β-sheet and the C-terminal kink-helix motif.

Discussion

Insights into the function from the 3D structure

Although it is clear that the dsRBP-ZFa protein binds to dsRNA and to RNA–DNA hybrids,[12,13] a specific function has not been identified for this protein. We have used the 3D structures described here to gain insight into possible RNA-binding sites for ZFa. The electrostatic potential of the dsRBP-ZFa zinc fingers is mapped onto the surface in Figure 6. A large positively charged face is formed by the helix-kink-helix motif in each of the fingers. The conserved basic residues at position +4 to the first histidine residue, R56 and K117, are solvent-exposed and can be expected to be involved in dsRNA recognition. The same holds true for R63/R124 at position +5 to the second histidine residue. Some degree of conservation is found at positions +2 and +3 relative to this histidine residue, although in the first fingers of dsRBP-ZFa one of these positions is occupied by an asparagine residue ($N_{61}K_{62}$ in finger I, $K_{121}N_{122}$ in finger II). Both lysine and asparagine are capable of forming hydrogen bonds to dsRNA. Thus, the structure suggests that the site of RNA binding is likely to be at the intersection between the two helices.

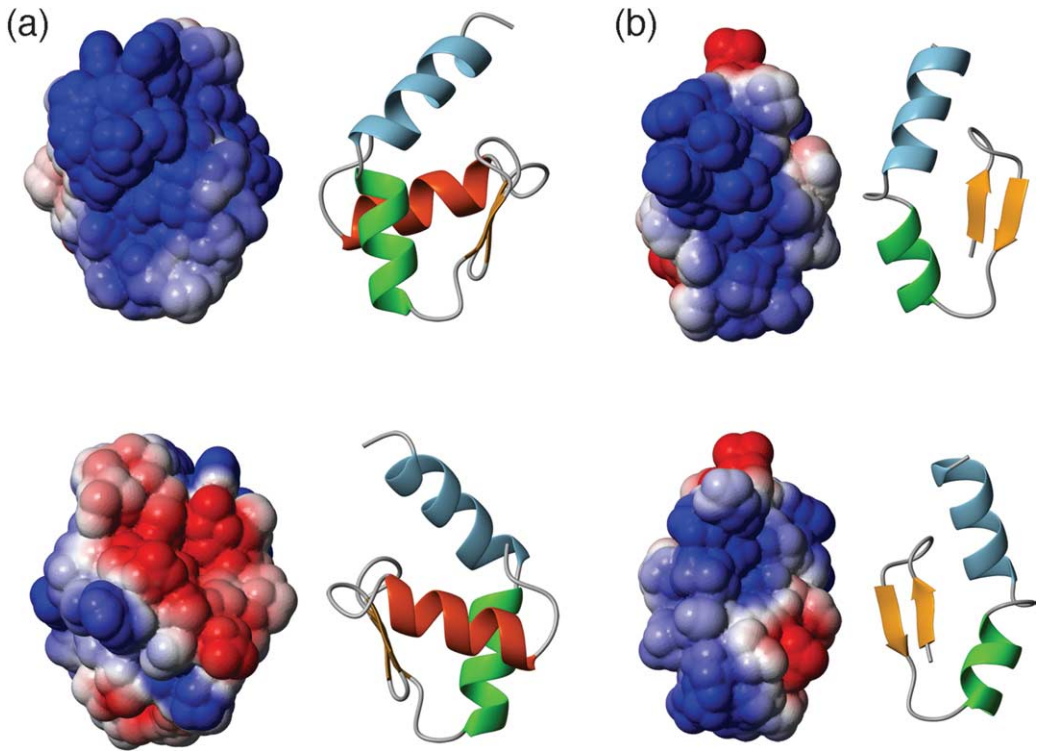

Figure 6. Surface electrostatic potential mapped in two orientations, rotated 180° between the upper and lower panels. (a) Finger I and (b) finger II from the lowest-energy structure. Blue corresponds to positive potential, white corresponds to neutral and red corresponds to negative potential. Intermediate values are colored using a linear interpolation between these colors. To the right of each surface representation are ribbon structures (colored according to the scheme of Figure 5) that show the orientations of the backbone in each case.

Structural basis for RNA *versus* DNA binding: comparison with TFIIIA and SWI5

The core secondary structure elements of the zinc finger domains of dsRBP-ZFa2-128 show some similarity with well-characterized dsDNA-binding canonical zinc finger folds. For example, finger II of ZFa shows a resemblance to the first three zinc fingers of TFIIIA[35] (Figure 7(a)), while finger I can be superimposed onto the backbone of the yeast zinc finger protein SWI5 (Figure 7(b)).[36]

Fitting of residues K95–C97 and C100–G116 of dsRBP-ZFa finger II onto Y13–C15 and C20–K36 of TFIIIA results in a backbone rmsd of 1.0 Å. The second β-strand, the turn into the α-helix and the first α-helix itself display a very good superposition, whereas the first β-strand and especially the turn between the strands show marked differences, as might be expected from the reduced spacing between the cysteine ligands of dsRBP-ZFa. The packing of side-chains in the protein cores is somewhat similar, in that the first three coordinating residues adopt similar positions and there is similar hydrophobic packing. Apart from these similarities, there are clear structural differences in terms of the coordination scheme, i.e. δ instead of ε coordination of the second histidine ligand of dsRBP-ZFa, and the presence of additional helical structure beyond this histidine ligand in the dsRNA-binding zinc fingers.

The similarity between ZFa finger I and the SWI5 zinc finger is limited to the overall topology and secondary structure content. If the zinc ligands are superimposed, some similarity can be seen between the second and third strand of the β-sheet and the second helix. In addition, the N-terminal helix of SWI5 is close to that of dsRBP-ZFa and points in roughly the same direction. Differences in the spacing between the zinc ligands gives rise to significant differences in backbone structure and turn geometry, and the β-sheet is more extensive in SWI5.

The TFIIIA zinc fingers 1–3 recognize dsDNA with a set of base-specific contacts in the major groove, supported by basic residues positioned to form salt-bridges with the phosphate backbone.[35] The TFIIIA zinc fingers interact with DNA through the turn leading into the α-helix and the helix itself, which occupy the major groove of the dsDNA. Although the ββα core of ZFa finger II could potentially fit into a dsDNA major groove, the side-chains that typically mediate zinc finger – DNA interactions in TFIIIA and other DNA-binding zinc finger proteins are missing from ZFa. Further, the positioning of the N-terminal helix (α_1) of finger 1 of ZFa would be expected to preclude binding of the fingertip and α_2 helix in the major groove of DNA. The weak affinity of ZFa for dsDNA may reflect the sequence differences that replace the characteristic DNA-binding side-chains, or may arise because

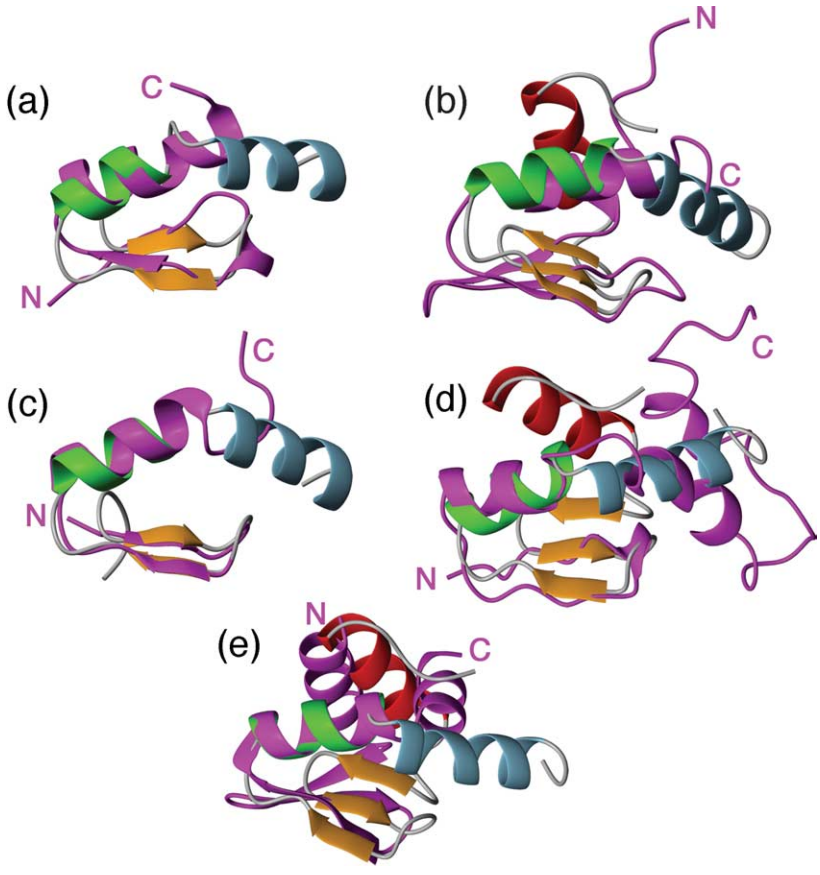

Figure 7. Superposition of (a) finger II of ZFa with finger 1 of TFIIIA (1tf3);[35] (b) finger I of ZFa with finger 1 of SWI5 (1ncs);[36] (c) finger II of ZFa with the first FOG zinc finger from U-shaped (1fv5);[55] (d) finger 1 of ZFa with the U1C zinc finger (1uw2);[38] (e) finger I of ZFa with the C-terminal domain of RecA (2reb).[40] In each case, the ZFa finger structures are colored according to the scheme in Figure 5, and the superimposed structures are colored magenta.

ZFa binds to DNA with a completely different binding mode from that of the canonical zinc finger.

Do the ZFa fingers participate in protein–protein interactions?

Superposition of the ββα core of finger II of ZFa with the first zinc finger of the FOG family protein U-shaped (USH)[37] (Figure 7(c)) shows an extremely high level of structural similarity between the two proteins (rmsd of 0.72 Å). The FOG fingers are not known as nucleic acid-binding proteins but interact with another type of zinc finger protein, GATA, *via* a hydrophobic face formed by residues in the second strand and in the helix. There is considerable similarity with the ZFa fingers in terms of hydrophobicity of this surface. It is thus possible that this face of the ZFa zinc finger structure is involved in interactions with other proteins, whereas nucleic acid interactions occur at the C-terminal kink-helix motif.

Similarity to the human spliceosomal protein U1C

The NMR structure of the spliceosomal zinc finger protein U1C[38] has a ββα zinc finger core that is C-terminally extended by a kink-helix motif and another short helix following an irregular turn. The ββα core shows great similarity to finger I of dsRBP-ZFa (Figure 7(d)), but the overall topology of the U1C finger is significantly different from that of the ZFa fingers. U1C does not contain the three-stranded β-sheet, and the C-terminal helix of U1C is parallel with the N-terminal helix of the first finger of dsRBP-ZFa, but shifted by about 5.5 Å perpendicular to the helix axis. The zinc coordination by the histidine residues of ZFa is in the less frequently observed ε,δ fashion. This leads to significant differences in the direct neighborhood of the coordination site and causes a completely different geometry of the kink between the helices, and affects the direction that the last helix of dsRBP-ZFa finger 1 adopts with respect to the protein core. When superimposed on the ββα core, the angle between the third helix of ZFa finger I and the second helix of U1C is about 30°.

Specific binding to RNA has been confirmed for U1C only when incorporated into the U1 snRNP.[38] However, the positively charged surface formed by the helix-kink-helix motif will likely be involved in RNA binding. The corresponding region of the dsRBP-ZFa fingers, particularly the kink and the following helix, is rich in basic amino acid residues and likely to form contacts with the phosphate backbone of dsRNA.

Despite these similarities, the ZFa zinc fingers are quite distinct from U1C and other Matrin-type zinc fingers, which contain conserved aromatic residues protruding from the β-sheet. Their function is still unclear, although RNA recognition through stacking interactions or hydrophobic interactions with other proteins has been discussed.[38] The dsRBP-ZFa fingers do not possess these aromatic residues, and the solvent-accessible surface of the β-sheet displays mostly hydrophilic residues. It is therefore likely that the ZFa-type zinc fingers are functionally distinct from those of the Matrin-type family.

Structural similarity with non-zinc-containing, RNA-binding proteins

A somewhat unexpected similarity was observed between the first zinc finger domain of dsRBP-ZFa and proteins that do not rely on zinc coordination in order to stabilize their tertiary structure. These include the C-terminal domain of the protein RecA from *Escherichia coli*,[39,40] which is involved in homologous recombination and DNA repair, and the well-known *Xenopus laevis* dsRNA-binding domain xlRBPa,[2,10] which recognizes RNA duplexes in a non-sequence-specific manner. The overall topology of the xlRBPa domain does not resemble the ZFa fingers at first glance, but there are similarities in certain structural elements. The N-terminal and the middle helix of finger I of ZFa can be aligned with the N-terminal and the end of the C-terminal helix of the xlRBPa (Figure 7(e)). This brings the three-stranded β-sheet of ZFa into the same plane as part of the much larger β-sheet of xlRBPa, although the strands do not line up but form an angle of about 45°. xlRBPa distinguishes dsRNA from dsDNA by probing the size of the major groove by bridging two adjacent minor grooves.[10,11] Such a mechanism would be impossible for ZFa as long as it adopts the structural alignment described above. First, the N-terminal α-helix is relatively poor in basic residues necessary to bind the phosphate backbone. Second, the β-sheet of ZFa would be too small to bridge the major groove and make contact with the next minor groove. We conclude that, despite the similarity between xlRBPa and the ZFa zinc fingers, the mechanism of RNA interaction with ZFa is most likely different from that of xlRBPa.

Finger I of dsRBP-ZFa shows an astonishing similarity to the C-terminal domain of the protein RecA from *E. coli*, which is involved in homologous recombination and DNA repair. This domain has all architectural elements of the αβββαα fold of ZFa, but it is not a zinc finger.[39,40] The domains can be superimposed with a backbone rmsd of 2.9 Å. The first two helices show good alignment, and the intervening β-sheets are aligned on one plane. The alignment of the remainder of the structures is worse, and the last helix points in a different direction, at an angle of almost 90°. Analogous to the non-sequence-specific recognition of dsRNA by dsRBP-ZFa, the C-terminal domain of RecA recognizes dsDNA non-sequence-specifically. The dsDNA-binding interface is located at the loop connecting the β-sheet and the middle helix, together with the first few residues of the helix itself and at the adjacent edge of the sheet.[39] However, the ZFa zinc finger structures show a much lower density of positively charged residues

in this region, and it therefore seems unlikely that this part of the structure constitutes the RNA-recognition surface, despite the topological similarity of the RecA structure.

Conclusions

The solution structure of the first two zinc finger domains of dsRBP-ZFa shows a number of features that suggest sites of RNA binding for this new zinc finger type. Although there are homologies with a number of other nucleic acid-binding protein domains, none of these homologous proteins gives a definitive picture that is applicable to the ZFa fingers. At the moment, on the basis of the electrostatic surfaces displayed in Figure 6, we suggest that there is probably a dsRNA-binding site located in the loop or kink between the two C-terminal helices of finger I, and that the same region of finger II could also accommodate RNA. The role of the linker sequence is not clear from the present studies, although the presence of positively charged side-chains suggests that it may participate in RNA binding. Although RNA binding by the ZFa zinc fingers is said to be non-specific,[41] we observe that not all RNA sequences appear to bind to the ZFa2-128 construct. Work is proceeding to identify an RNA sequence with a sufficiently high binding affinity that 3D structures of the complex can be calculated.

Experimental Procedures

Cloning

Fragments corresponding to residues 1–74, 1–128 and 1–205 of dsRNAbp ZFa were amplified by PCR from a *Xenopus* blastula cDNA library (kindly donated by Joel Gottesfeld) and cloned into the phage T7 expression vector pET21a (Novagen). Expression of soluble protein was observed for all constructs upon addition of zinc to the culture medium. In the case of ZFa 1-128 and 1-205, an 80 residue truncation product was expressed in addition to the full-length proteins. The appearance of this byproduct was eliminated through the introduction of silent mutations at codons 77 and 78.

Expression and purification

Expression of isotopically labeled dsRBP-ZFa2-128 was carried out at 37 °C in M9-minimal medium supplemented with $^{15}NH_4Cl$, $(^{15}NH_4)_2SO_4$ and [^{13}C]glucose. Zinc was added upon induction to 150 μM. When the cell density reached a plateau, usually after about six hours, cells were harvested by centrifugation, resuspended in 20 mM Mes (pH 6.1), 10 mM DTT, 50 μM $ZnSO_4$, and lysed by sonication. The protein was purified from the soluble fraction under native conditions at 4 °C by FPLC using a combination of cation-exchange and anion-exchange columns followed by gel-filtration chromatography. The purity of the protein was checked by SDS-PAGE and matrix-assisted laser desorption/ionization mass spectrometry.

CD spectroscopy

Ultraviolet circular dichroism spectra were collected using an Aviv model 202 CD spectrometer at 25 °C, using a 0.2 cm cell. Typical protein concentration was 10 μM.

NMR sample preparation

The protein was transferred into NMR sample buffer (25 mM Tris (pH 6.9), 100 mM NaCl, 2 mM DTT) by ultracentrifugation using Millipore 15 mp Centriprep YM3 and 2 mp Centricon YM3 filter devices with a cut-off of 3000 Da. At the same time, the concentration was increased to ca 1 mM. For experiments in ^2H_2O, the sample was repeatedly diluted in NMR sample buffer made with ^2H_2O and re-concentrated.

NMR spectroscopy

NMR spectra were recorded on Bruker DRX600, DMX750 or Avance900 spectrometers at 298 K, processed using NMRPipe,[42] and analyzed using NMRView.[43] Backbone resonances were assigned from 3D HNCACB,[44] CBCA(CO)NH,[45] HBHA(CBCACO)NH,[46] and ^{15}N-resolved total correlated spectroscopy (TOCSY)-HSQC and NOESY-HSQC spectra.[47] (H)CCH-COSY and (H)CCH-TOCSY experiments[48,49] were performed; nearly complete side-chain resonance assignments were obtained for dsRBP-ZFa2-128. The backbone amide protons of K57 and S105 were not observed. The β-protons of 18 amino acid residues were assigned stereospecifically using coupling constants derived from the 3D HNHA,[20] and HACAHB-COSY[50] experiments, and relative NOE intensities. Stereospecific assignment for valine γ-methyl groups was accomplished by ^{13}C–[^{13}CO] and ^{13}C–[^{15}N] spin-echo difference CT-HSQC spectra.[15] The ^{1}H–^{15}N steady-state heteronuclear NOEs were measured at 14.1 T for backbone amide groups using water flip-back methods.[51] Zinc coordinating histidine residues were identified from ^{1}H–^{15}N-HMQC spectra,[30] which revealed that the coordination occurs through $N^{\varepsilon 2}$ of H52 and H113 and through $N^{\delta 1}$ of H58 and H119.

Restraint generation, structure calculation and analysis

Inter-proton upper bound distance restraints were derived from crosspeak volumes in 3D ^{15}N-separated NOESY-HSQC at 600 MHz and 3D ^{13}C-separated NOESY-HSQC at 900 MHz. Initial distance restraints were obtained with the CANDID module[19] of the program CYANA.[18] Separate chemical shift lists were used for the two NOESY spectra to account for slight variations in sample conditions. Initial CANDID runs did not involve any manually added restraints or restraints that would support a specific coordination scheme for zinc. A list of ligand–ligand distances obtained from structures calculated in the absence of metal-ligand constraints, and an HMQC spectrum showing the histidine crosspeaks upon which we based our ε/δ coordination assignments are included in the Supplementary Data. In later stages, restraints for zinc coordination and restraints supporting confirmed elements of secondary structure were provided in order to optimize the yield of automatically assigned NOE connectivities. CYANA calculations were performed using ten processors of a 128-processor Silicon Graphics computer in parallel.

Backbone ϕ, ψ and side-chain χ_1 torsion angle restraints were derived from ^3J coupling constants measured by 3D HNHA, 3D HACAHB-COSY, $^{13}C[^{13}CO]$ and $^{13}C[^{15}N]$ spin-echo difference CT-HSQC spectra,[20,22,23,50] in combination with chemical shift data analyzed with the TALOS program,[21] and the pattern of medium-range NOE connectivities.

The final set of restraints was obtained by manually resolving ambiguities and adjusting upper distance bounds to account for spectral overlap and relaxation behavior. A total of 1142 unique NOE restraints and 216 dihedral angle restraints were used in the structure calculations (Table 1). The refined set of restraints was used by CYANA[18] to yield an initial set of 400 structures. The 100 structures with lowest target function were imported into AMBER 8 using a PARM94 force field modified to reduce charges to 20% of their original values, 20 kcal/(mol Å^2) force constant on inter-proton distance restraints and 150 kcal/(mol rad^2) for torsion angle restraints. An additional set of 325 amino acid chirality restraints and 127 restraints restricting the peptide bond to the *trans* configuration for the majority of the residues was active during AMBER calculations. Zinc–sulfur and zinc–nitrogen bond lengths were set to 2.3 Å and 2.1 Å, respectively, and the coordination geometry was constrained to be tetrahedral.

After five cycles of simulated annealing and final energy minimization the 50 lowest-energy structures were transferred to a full-charge force field (PARM99) and subjected to another five cycles of simulated annealing using a generalized Born solvent model.[25] All AMBER calculations were performed on a 512 processor Linux cluster using one Intel XEON processor per structure.

A total of 20 structures with lowest combined distance and angle violations were sorted by restraint violation energy and selected for analysis. Backbone analysis was performed using PROCHECK[52] and PROMOTIF.[53] Graphics images were prepared with MOLMOL[54] and POV-Ray (©POV-Team™ 1991–2002).

Data Bank accession numbers

The coordinates and NMR restraint files have been deposited in the RCSB Protein Data Bank (1ZU1) and the chemical shift assignments in the BioMagResBank (6655).

Acknowledgements

We thank Peter Güntert and Wolfgang Peti for assistance with early versions of CYANA, Gerard Kroon and John Chung for assistance with NMR experiments, Brian Lee for helpful discussions, and Linda Tennant for expert technical assistance. This work was supported by grant GM36643 from the National Institutes of Health and a DFG post-doctoral research fellowship to H.M.

References

1. Nicholson, A. W. (1996). Structure, reactivity, and biology of double-stranded RNA. *Prog. Nucl. Acid Res. Mol. Biol.* **52**, 1–65.
2. Saunders, L. R. & Barber, G. N. (2003). The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB J.* **17**, 961–983.
3. Fierro-Monti, I. & Mathews, M. B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem. Sci.* **25**, 241–246.
4. Bernstein, E., Denli, A. M. & Hannon, G. J. (2001). The rest is silence. *RNA*, **7**, 1509–1521.
5. Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer, New York.
6. Carlson, C. B., Stephens, O. M. & Beal, P. A. (2003). Recognition of double-stranded RNA by proteins and small molecules. *Biopolymers*, **70**, 86–102.
7. Batey, R. T., Rambo, R. P. & Doudna, J. A. (1999). Tertiary motifs in RNA structure and folding. *Angew. Chem. (Engl)*, **38**, 2327–2343.
8. Hermann, T. & Patel, D. J. (1999). Stitching together RNA tertiary architectures. *J. Mol. Biol.* **294**, 829–849.
9. Stefl, R., Skrisovska, L. & Allain, F. H. (2005). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* **6**, 33–38.
10. Ryter, J. M. & Schultz, S. C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **17**, 7505–7513.
11. Ramos, A., Bayer, P. & Varani, G. (1999). Determination of the structure of the RNA complex of a double-stranded RNA-binding domain from Drosophila Staufen protein. *Biopolymers*, **52**, 181–196.
12. Finerty, P. J. & Bass, B. L. (1997). A *Xenopus* zinc finger protein that specifically binds dsRNA and RNA–DNA hybrids. *J. Mol. Biol.* **271**, 195–208.
13. Finerty, P. J. & Bass, B. L. (1999). Subsets of the zinc finger motifs in dsRBP-ZFa can bind double-stranded RNA. *Biochemistry*, **38**, 4001–4007.
14. Laity, J. H., Dyson, H. J. & Wright, P. E. (2000). DNA-induced α-helix capping in conserved linker sequences is a determinant of binding affinity in Cys_2-His_2 zinc fingers. *J. Mol. Biol.* **295**, 719–727.
15. Varmeh-Ziaie, S., Okan, I., Wang, Y. S., Magnusson, K. P., Warthoe, P., Strauss, M. & Wiman, K. G. (1997). Wig-1, a new p53-induced gene encoding a zinc finger protein. *Oncogene*, **15**, 2699–2704.
16. Israeli, D., Tessler, E., Haupt, Y., Elkeles, A., Wilder, S., Amson, R. *et al*. (1997). A novel p53-inducible gene, PAG608, encodes a nuclear zinc finger protein whose overexpression promotes apoptosis. *EMBO J.* **16**, 4384–4392.
17. Yang, M., May, W. S. & Ito, T. (1999). JAZ requires the double-stranded RNA-binding zinc finger motifs for nuclear localization. *J. Biol. Chem.* **274**, 27399–27406.
18. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
19. Herrmann, T., Güntert, P. & Wüthrich, 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.
20. Vuister, G. W. & 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.
21. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
22. Grzesiek, S., Vuister, G. W. & 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.
23. Vuister, G. W., Wang, A. C. & 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.
24. Case, D. A., Cheatham, T., Darden, T., Gohlke, H., Luo, R. & Merz, K. M. J. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* in the press.
25. Bashford, D. & Case, D. A. (2000). Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **51**, 129–152.
26. Xia, B., Tsui, V., Case, D. A., Dyson, H. J. & 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.
27. Ludvigsen, S. & Poulsen, F. M. (1992). Positive ϕ-angles in proteins by nuclear magnetic resonance spectroscopy. *J. Biomol. NMR*, **2**, 227–233.
28. Gronenborn, A. M. & Clore, G. M. (1994). Identification of N-terminal helix capping boxes by means of ^{13}C chemical shifts. *J. Biomol. NMR*, **4**, 455–458.
29. Pascual, J., Martinez-Yamout, M., Dyson, H. J. & Wright, P. E. (2000). Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. *J. Mol. Biol.* **304**, 723–729.
30. Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. (1993). Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated III^{Glc}, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci.* **2**, 543–558.
31. Holm, L. & Sander, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480.
32. Krissinel, E. & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallog. sect. D*, **60**, 2256–2268.
33. Michalopoulos, I., Torrance, G. M., Gilbert, D. R. & Westhead, D. R. (2004). TOPS: an enhanced database of protein structural topology. *Nucl. Acids Res.* **32**, D251–D254.
34. Gibrat, J. F., Madej, T. & Bryant, S. H. (1996). Surprising similarities in structure comparison. *Curr. Opin. Struct. Biol.* **6**, 377–385.
35. Wuttke, D. S., Foster, M. P., Case, D. A., Gottesfeld, J. M. & Wright, P. E. (1997). Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. *J. Mol. Biol.* **273**, 183–206.
36. Dutnall, R. N., Neuhaus, D. & Rhodes, D. (1996). The solution structure of the first zinc finger domain of SW15: a novel structural extension to a common fold. *Structure*, **4**, 599–611.
37. Liew, C. K., Kowalski, K., Fox, A. H., Newton, A., Sharpe, B. K., Crossley, M. & Mackay, J. P. (2000). Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that mediate protein–protein interactions. *Struct. Fold. Des.* **8**, 1157–1166.
38. Muto, Y., Krummel, D. P., Oubridge, C., Hernandez, H., Robinson, C. V., Neuhaus, D. & Nagai, K. (2004). The structure and biochemical properties of the human spliceosomal protein U1C. *J. Mol. Biol.* **341**, 185–198.
39. Aihara, H., Ito, Y., Kurumizaka, H., Terada, T., Yokoyama, S. & Shibata, T. (1997). An interaction between a specified surface of the C-terminal domain of RecA protein and double-stranded DNA for homologous pairing. *J. Mol. Biol.* **274**, 213–221.
40. Story, R. M., Weber, I. T. & Steitz, T. A. (1992). The structure of the *E. coli recA* protein monomer and polymer. *Nature*, **355**, 318–325.
41. Bass, B. L., Hurst, S. R. & Singer, J. D. (1994). Binding properties of newly identified *Xenopus* proteins containing dsRNA-binding motifs. *Curr. Biol.* **4**, 301–314.
42. Delaglio, F., Grzesiek, S., Vuister, G. W., Guang, Z., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
43. Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 604–613.
44. Wittekind, M. & 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.
45. Grzesiek, S. & 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.
46. Grzesiek, S. & 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.
47. Fesik, S. W. & 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.
48. Bax, A., Clore, G. M. & Gronenborn, A. M. (1990). ^{1}H–^{1}H correlation *via* isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson.* **88**, 425–431.
49. Gehring, K. & Ekiel, I. (1998). H(C)CH-COSY and (H) CCH-COSY experiments for ^{13}C-labeled proteins in H_2O solution. *J. Magn. Reson.* **135**, 185–193.
50. Grzesiek, S., Kuboniwa, H., Hinck, A. P. & 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.
51. Grzesiek, S. & 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.
52. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & 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.
53. Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF-A program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.
54. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
55. Liew, C. K., Simpson, R. J., Kwan, A. H., Crofts, L. A., Loughlin, F. E., Matthews, J. M. *et al*. (2005). Zinc fingers as protein recognition motifs: structural basis for the GATA-1/friend of GATA interaction. *Proc. Natl Acad. Sci. USA*, **102**, 583–588.

Edited by M. F. Summers