Full text provided by www.sciencedirect.com

SCIENCE DIRECT®

Zinc finger proteins: getting a grip on RNA

Raymond S Brown

C2H2 (Cys-Cys-His-His motif) zinc finger proteins are members of a large superfamily of nucleic-acid-binding proteins in eukaryotes. On the basis of NMR and X-ray structures, we know that DNA sequence recognition involves a short α helix bound to the major groove. Exactly how some zinc finger proteins bind to double-stranded RNA has been a complete mystery for over two decades. This has been resolved by the long-awaited crystal structure of part of the TFIIIA–5S RNA complex. A comparison can be made with identical fingers in a TFIIIA–DNA structure. Additionally, the NMR structure of TIS11d bound to an AU-rich element reveals the molecular details of the interaction between CCCH fingers and single-stranded RNA. Together, these results contrast the different ways that zinc finger proteins bind with high specificity to their RNA targets.

Addresses
Division of Experimental Medicine, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine, 4 Blackfan Circle, Boston, MA 02115, USA

Corresponding author: Brown, Raymond S (rbrown7@bidmc.harvard.edu)

Current Opinion in Structural Biology 2005, **15**:94–98

This review comes from a themed issue on
Protein–nucleic acid interactions
Edited by Aneel K Aggarwal and Jennifer A Doudna

Available online 26th January 2005

0959-440X/$ – see front matter
© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.sbi.2005.01.006

Introduction

Many DNA-binding proteins have multiple copies of small independently folded domains that contain conserved cysteines and histidines coordinated to zinc; such proteins are commonly called zinc finger proteins. Several different types of Cys-Cys (CC) or Cys-His (CH) motifs are present in zinc finger proteins from a wide variety of organisms [1••]. In the past few years, we have seen a steady trickle of reports on zinc finger proteins that show RNA binding activity. These include viral proteins, such as the HIV-1 nucleocapsid (CCHC) [2,3], reovirus σ3 (C2H2) [4] and barley stripe mosaic virus γb protein (C4C/H) [5]. Examples also come from a plant, *Arabidopsis* HUA1 nuclear protein (CCCH) [6], and parasites, such as trypanosome tcZFP1 (CCCH) [7] and leishmania mitochondrial RET1 uridylyl transferase (C2H2) [8]. The mammalian zinc finger proteins wig-1 (C2H2) [9] and JAZ (C2H2) [10] are localized to the nucleolus, whereas others, such as hZFP100 (C2H2) [11] and tristetraprolin TTP (CCCH) [12], are involved in histone pre-mRNA processing and the degradation of tumor necrosis factor α mRNA, respectively. In addition, there are reports of dual RNA/DNA-binding proteins, such as the thyroid hormone receptor (CCCC) [13] and the trypanosome poly-zinc finger PZFP1 pre-mRNA processing protein (CCHC) [14]. Whether their interactions with RNA are based on the same mechanisms as protein–DNA binding is an intriguing structural question that has remained unanswered until now.

What follows is an attempt to expose both similarities and differences between C2H2 zinc finger protein binding to RNA and DNA based on recent X-ray structures. Examples of ssRNA and dsRNA sequence recognition in are also examined and discussed at the molecular level.

DNA sequence recognition

The structural details of how zinc fingers bind to dsDNA are well understood. Zinc fingers typically follow a right-handed helical path as they wrap around the outside of a double helix. Several protein–DNA complexes show that multiple contacts are made in particular with nucleotide bases in the major groove [15]. A single C2H2 zinc finger, composed of a β hairpin and an α helix held together by a tetrahedrally coordinated zinc ion, will span a DNA sequence of three or four consecutive base pairs. Frequently, the contacts are made by the sidechains of amino acids located at positions −1, +2, +3 and +6 of the α helix. Strong preferences are observed, for example, arginine bonding to guanosine, aspartic acid to adenosine or cytidine, and leucine to thymidine. At present, these observations are insufficient to define a set of coding rules.

TFIIIA and DNA

TFIIIA is necessary for the transcription by RNA polymerase III of the genes encoding eukaryotic ribosomal 5S RNA. It is the essential core component of the initiation complex required for the start of transcription. We know that not all of the nine C2H2 fingers in TFIIIA bind to DNA base pairs. Some have a passive role as spacers that allow the protein to span a long DNA sequence [16]. In this way, TFIIIA interacts with the separate transcription signals located within the 55 base pairs of the ribosomal 5S RNA gene internal promoter region.

TFIIIA also binds to 5S RNA

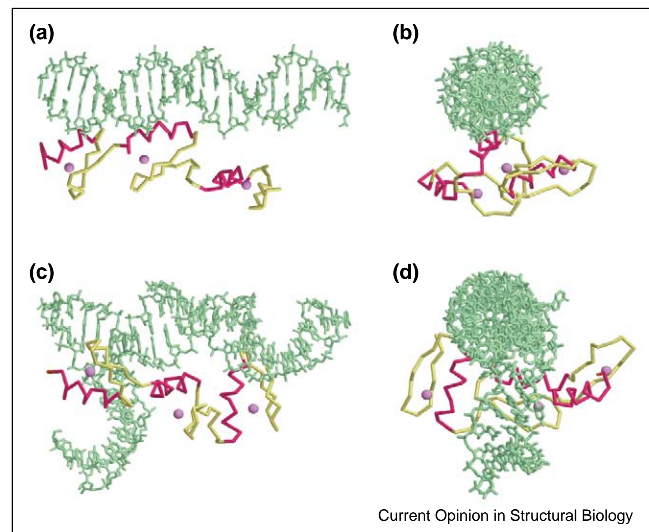
A large amount of TFIIIA protein is found associated with ribosomal 5S RNA in cytoplasmic 7S storage particles from the oocytes of amphibians and spiny fish. How can TFIIIA use its zinc fingers to bind specifically to DNA and RNA, two quite distinct forms of the double

helix with different conformations? What kind of structural arrangement would allow zinc fingers to bind to both? This fascinating question has occupied many people for more than 20 years. Early suggestions were made involving DNA-like structures in the internal loops of 5S RNA or an intermediate double-helical conformation of 5S DNA. Because base pairs in the major groove of RNA are relatively inaccessible, recognition may involve surface contacts, especially to negatively charged phosphates. It is now generally believed that subsets of zinc fingers have adapted either to RNA binding or to DNA binding. Thus, fingers 1–3 of TFIIIA bind strongly to DNA and central fingers 4–6 bind to RNA [17].

Divide and conquer

Large single crystals of the *Xenopus* oocyte 7S complex diffract only to about 8 Å resolution (RS Brown, E Lorentzen, FA Rey, unpublished). Lu *et al.* [18••] have achieved a remarkable breakthrough by engineering a smaller protein–RNA complex. They used a 'divide-and-conquer' strategy [19] to custom design a truncated 5S RNA that binds to zinc fingers 4–6 of TFIIIA. They obtained crystals that diffract to 3 Å and used them to successfully determine the X-ray structure. Their structure reveals the details of how fingers 4–6, which contribute the most to RNA recognition, interact with a 61-nucleotide RNA that correctly mimics *Xenopus* oocyte 5S RNA, comprising loops A and E, as well as helix V and a shortened helix IV. Finally we can admire the view and see how a highly adapted DNA-binding zinc finger protein interacts with a folded RNA (Figure 1).

Figure 1



Dual nucleic acid binding by zinc fingers of TFIIIA. **(a)** Interaction of fingers 4–6, from right to left, with 5S DNA [16]. **(b)** Endwise view of (a). **(c)** Interaction of fingers 4–6, from right to left, with a truncated 5S RNA [18••]. **(d)** Endwise view of (c). Nucleic acids are shown in green. Fingers 4–6 are represented as a backbone trace, with their α helices in red and zinc ions as violet spheres. Figure generated using the program RasMol [28].

Similar but different

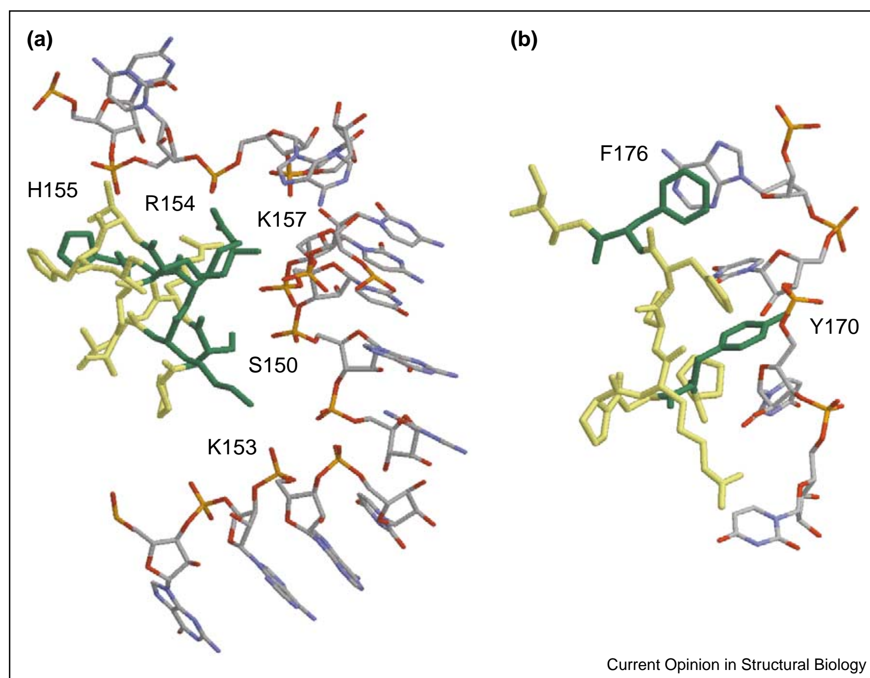
From these crystal structures, we can now see that fingers 4–6 are associated with identical sequences in their respective RNA and DNA complexes. Surprisingly, the α helices of fingers 4–6 are largely responsible for the interaction with RNA, but are excluded from its deep and narrow major groove. As expected, there are only a few contacts to nucleotide bases, including two of the triple base pairs found at junctions between helical stems and internal loops in 5S RNA. His119 of finger 4 binds to bulged G75 (numbering refers to the intact 5S RNA) and possibly also to G99 in internal loop E. Finger 5 has no direct interactions with the bases, whereas Trp177 of finger 6 stacks onto A11 in loop A. This resembles the hydrophobic stacking of Trp37 from the HIV-1 nucleocapsid (CCHC box) second zinc finger domain with a guanosine base in the terminal loop of the SL3 RNA hairpin [20]. Curiously, some residues in finger 5 (e.g. Ser150, Lys144, Arg154 and His155) bind RNA and DNA, but to different sites in each case (Figure 2a).

It is quite remarkable that the relative orientation of fingers 4 and 5 remains the same regardless of whether the fingers are bound to RNA or DNA. However, the α helix of finger 4 now points into the major groove of loop E, close to phosphate 76, in complete contrast to its role as a passive spacer over the minor groove of DNA. Surprisingly, finger 5 has tightened its grip on the RNA and contacts phosphates in helices 1 and 3. The α helix of finger 5 lies across the major groove, binding to phosphates 68, 69 and 70 on one strand and to phosphate 101 on the other. However, in the complex with DNA, this helix follows the major groove and interacts with nucleotide bases G70 and G71. Thus, if fingers 4 and 5 of both structures are superimposed, the RNA and DNA double helices are orientated roughly perpendicular to each other. Oddly, finger 6, which spans a minor groove of the DNA, has rotated around its short Ala-Gly linker connecting fingers 5 and 6 to interact with loop A and phosphates 10 and 11 in the RNA complex. It is premature to rationalize the observed flexibility of finger 6, as its correct positioning may actually require additional fingers 7–9 to be present.

Zinc fingers and single-stranded RNA

The trafficking, activity and stability of eukaryotic mRNAs are mediated by proteins containing a single or combinations of distinct RNA-binding domains. Some of these proteins recognize and bind to AU-rich sequences in the 3′-untranslated regions of mRNA. Two copies of a CCCH finger domain are present in the immediate early response gene products Nup475 and TIS11d; these proteins bind to AU-rich elements of tumor necrosis factor α mRNA, resulting in the destabilization

Figure 2



Molecular mechanisms of ssRNA and dsRNA recognition by zinc finger domains. **(a)** C2H2 finger 5 of TFIIIA makes contact with phosphates in 5S dsRNA stems I and V [18••] using the sidechains of α-helix residues Ser150, Lys153, Arg154, Lys157 and His155 (green). **(b)** Novel intercalation between ssRNA bases by conserved hydrophobic residues (green) in the 3_{10}-helix region of TIS11d CCCH zinc fingers [23••]. High-affinity binding is dependent on aromatic stacks U8-Tyr170-U9 and U6-Phe176-A7.

of the mRNA [21]. The NMR structure of a single CCCH domain from Nup475 reveals the presence of a short α helix between the first and second cysteines that are coordinated to zinc, but little or no other secondary structure is present [22]. Wright and co-workers [23••] have reported the first NMR structure of a complex between tandem CCCH domains of TIS11d and the ssRNA sequence 5′-UUAUUUAUU-3′. Surprisingly, the zinc finger domains of Nup475 and TIS11d are folded differently and have differing metal coordination. Perhaps this discrepancy might be the result of unfavorable conditions used to refold the recombinant proteins?

Recognition of the AU-rich element

As expected, each CCCH domain of TIS11d binds separately to its own 5′-UAUU-3′ half-site. The predominant interaction between the protein and ssRNA is the intercalation of a tyrosine sidechain between a UU dinucleotide and a phenylalanine sidechain between the following AU dinucleotide. This hydrophobic stacking of aromatic rings and heterocyclic bases, and their locations provide the basis of sequence recognition (Figure 2b). These interactions interrupt the normal nucleotide base stacking and the RNA chain is consequently kinked at each adenosine position. The two finger domains squat neatly above the exposed nucleotide bases and the intervening 18 amino acid linker sequence runs parallel to the phosphodiester backbone. No contacts are described between amino acids and the phosphates. There are additional stabilizing hydrogen bonds between several of the conserved N-terminal K/RYKTEL sequence mainchain amides and carbonyl groups and the Watson–Crick edges of the bases.

RNA-binding zinc finger proteins

Some NMR structures of zinc finger proteins without their RNA targets are worthy of brief mention. Archaeal 30S ribosomal protein S27e contains a C4 (CCCC) finger motif whose structure was shown by NMR to consist of a β sandwich comprising two three-stranded β sheets [24]. The four metal-binding cysteine residues are located in the proximal loops. Eukaryotic S27 proteins are involved in the processing of damaged mRNA.

The U1C protein is a component of the U1 small nuclear ribonucleoprotein particle of the spliceosome and recognizes the 5′ splice site of pre-mRNA. The NMR structure reveals a TFIIIA-type C2H2 zinc finger, except that the canonical α helix is interrupted by a five-residue intervening loop between the two zinc-coordinating histidine residues [25]. The finger domain of U1C is extended by two C-terminal α helices.

TLS is involved in the nuclear export of spliced mRNA and contains several motifs involved in RNA binding, namely RGG (Arg-Gly-Gly) repeats, several RNA recognition motif (RRM) domains and a C4 zinc finger domain [26]. The structure of the C4 finger domain was solved by NMR and homology model building. It has a β hairpin-loop fold and cysteine residues coordinating zinc, and is shown to bind to GGUG-containing RNA. TLS has 44% sequence homology with the ZNF265 protein, which binds to cyclin mRNA [27].

Conclusions

Our first glimpse into the world of RNA recognition by zinc finger proteins reveals the existence of many different kinds of amino acid–nucleotide interactions. The mechanism of RNA sequence recognition is unlike the mechanism commonly used for the recognition of dsDNA, whereby a short α helix hydrogen bonds to major groove bases. Binding to dsRNA depends heavily on contacts with phosphates and there are also hydrophobic stacking interactions with accessible nucleotide bases that are fortuitously located to provide unique sequence specificity. In TFIIIA, a dual RNA/DNA-binding protein, we may have a case of 'one structure fits all', whereby individual C2H2 fingers make different contacts with each of the nucleic acids. ssRNA binding involves the intercalation of protein aromatic sidechains between appropriately spaced dinucleotide bases. This is supplemented by protein mainchain amide and carbonyl hydrogen bonds to the bases. Clearly, the specificity of the ssRNA interaction can be enhanced through repetition by the addition of a second CCCH finger domain.

The determination of a new RNA–protein structure is still a comparatively rare and exotic event. Our knowledge of these kinds of molecular interactions has been increased by recent success with the prokaryotic ribosome ([24] and references therein). Understanding the basis of protein–RNA interactions in splicing, nuclear transport, interference, translation and virus replication depends largely on increasing our efforts.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Iuchi S, Kudell N (Eds): *Zinc Finger Proteins: From Atomic Contact* •• *to Cellular Function*. Georgetown, TX: Landes Bioscience; 2004.
This book completely covers all aspects of the different classes of zinc finger proteins and their cellular activities.

2. Tisne C, Roques BP, Dardel F: **Specific recognition of primer tRNALys3 by HIV-1 nucleocapsid protein: involvement of the zinc fingers and the N-terminal basic extension**. *Biochimie* 2003, **85**:557-561.

3. Ramboarina S, Druillennec S, Morellet N, Bouaziz S, Roques BP: **Target specificity of human immunodeficiency virus type 1 NCp7 requires an intact conformation of its CCHC N-terminal zinc finger**. *J Virol* 2004, **78**:6682-6687.

4. Olland AM, Jane-Valbuena J, Schiff LA, Nibert ML, Harrison SC: **Structure of the reovirus outer capsid and dsRNA-binding protein sigma3 at 1.8 Å resolution**. *EMBO J* 2001, **20**:979-989.

5. Bragg JN, Lawrence DM, Jackson AO: **The N-terminal 85 amino acids of the barley stripe mosaic virus gammab pathogenesis protein contain three zinc-binding motifs**. *J Virol* 2004, **78**:7379-7391.

6. Cheng Y, Kato N, Wang W, Li J, Chen X: **Two RNA-binding proteins, HEN4 and HUA1, act in the processing of AGAMOUS pre-mRNA in *Arabidopsis thaliana***. *Dev Cell* 2003, **4**:53-66.

7. Espinosa JM, Portal D, Lobo GS, Pereira CA, Alonso GD, Gomez EB, Lan GH, Pomar RV, Flawia MM, Torres HN: ***Trypanosoma cruzi* poly-zinc finger protein: a novel DNA/RNA-binding CCHC-zinc finger protein**. *Mol Biochem Parasitol* 2003, **131**:35-44.

8. Aphasizheva I, Aphasizhev R, Simpson L: **RNA-editing terminal uridylyl transferase 1: identification of functional domains by mutational analysis**. *J Biol Chem* 2004, **279**:24123-24130.

9. Mendez-Vidal C, Wilhelm MT, Hellborg F, Qian W, Wiman KG: **The p53-induced mouse zinc finger protein wig-1 binds double-stranded RNA with high affinity**. *Nucleic Acids Res* 2002, **30**:1991-1996.

10. Chen T, Brownawell AM, Macara IG: **Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5**. *Mol Cell Biol* 2004, **24**:6608-6619.

11. Dominski Z, Erkmann JA, Yang X, Sanchez R, Marzluff WF: **A novel zinc finger protein is associated with U7 snRNP and interacts with the stem-loop binding protein in the histone pre-mRNP to stimulate 3′ end processing**. *Genes Dev* 2002, **16**:58-71.

12. Brewer BY, Malicka J, Blackshear PJ, Wilson GM: **RNA sequence elements required for high affinity binding by the zinc finger domain of tristetraprolin: conformational changes coupled to the bipartite nature of AU-rich mRNA-destabilizing motifs**. *J Biol Chem* 2004, **279**:27870-27877.

13. Xu B, Koenig RJ: **An RNA-binding domain in the thyroid hormone receptor enhances transcriptional activation**. *J Biol Chem* 2004, **279**:33051-33056.

14. Morking PA, Dallagiovanna BM, Foti L, Garat B, Picchi GF, Umaki AC, Probst CM, Krieger MA, Goldenberg S, Fragoso SP: **TcZFP1: a CCCH zinc finger protein of *Trypanosoma cruzi* that binds poly-C oligoribonucleotides *in vitro***. *Biochem Biophys Res Commun* 2004, **319**:169-177.

15. Wolfe SA, Nekludova L, Pabo CO: **DNA recognition by Cys2His2 zinc finger proteins**. *Annu Rev Biophys Biomol Struct* 2000, **29**:183-212.

16. Nolte RT, Conlin RM, Harrison SC, Brown RS: **Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex**. *Proc Natl Acad Sci USA* 1998, **95**:2938-2943.

17. Searles MA, Lu D, Klug A: **The role of the central zinc fingers of transcription factor IIIA in binding to 5S RNA**. *J Mol Biol* 2000, **301**:47-60.

18. Lu D, Searles A, Klug A: **Crystal structure of a zinc-finger-RNA** •• **complex reveals two modes of molecular recognition**. *Nature* 2003, **426**:96-100.
The first successful X-ray crystal structure of a classical C2H2 zinc finger–RNA complex to be described.

19. Berg JM: **Fingering nucleic acids: the RNA did it**. *Nat Struct Biol* 2003, **10**:986-987.

20. De Guzman RN, Wu ZR, Stalling CC, Pappalardo L, Borer PN, Summers MF: **Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element**. *Science* 1998, **279**:384-388.

21. Blackshear PJ: **Tristetraprolin and other CCCH tandem zinc finger proteins in the regulation of mRNA turnover**. *Biochem Soc Trans* 2002, **30**:945-952.

22. Amann BT, Worthington MT, Berg JM: **A Cys3His zinc-binding domain from Nup475/tristetraprolin: a novel fold with a disklike structure**. *Biochemistry* 2003, **42**:217-221.

23. Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE:
•• **Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d**. *Nat Struct Mol Biol* 2004, **11**:257-264.
This paper shows the NMR structure of two zinc finger domains binding in tandem to nine residues of ssRNA using a novel mechanism of hydrophobic intercalation.

24. Herve Du Penhoat C, Atreya HS, Shen Y, Liu G, Acton TB, Xiao R, Li Z, Murray D, Montelione GT, Szyperski T: **The NMR solution of the 30S ribosomal protein S27e encoded in gene RS27_ARCFU of *Archaeoglobus fulgidis* reveals a novel protein fold**. *Protein Sci* 2004, **13**:1407-1416.

25. Muto Y, Pomeranz Krummel D, Oubridge C, Hernandez H, Robinson CV, Neuhaus D, Nagai K: **The structure and biochemical properties of the human spliceosomal protein U1C**. *J Mol Biol* 2004, **341**:185-198.

26. Iko Y, Kodama TS, Kasai N, Oyama T, Morita EH, Muto T, Okumura M, Fujii R, Takumi T, Tate S, Morikawa K: **Domain architectures and characterization of an RNA binding protein, TLS**. *J Biol Chem* 2004, **279**:44834-44840.

27. Plambeck CA, Kwan AHY, Adams DJ, Westman BJ, van der Weyden L, Medcalf RL, Morris BJ, Mackay JP: **The structure of the zinc finger domain from human splicing factor ZNF265 fold**. *J Biol Chem* 2003, **278**:22805-22811.

28. Bernstein HJ: **Recent changes to RasMol, recombining the variants**. *Trends Biochem Sci* 2000, **25**:453-455.