Zinc finger proteins: new insights into structural and functional diversity

John H Laity, Brian M Lee and Peter E Wright*

Zinc finger proteins are among the most abundant proteins in eukaryotic genomes. Their functions are extraordinarily diverse and include DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding. Zinc finger structures are as diverse as their functions. Structures have recently been reported for many new zinc finger domains with novel topologies, providing important insights into structure/function relationships. In addition, new structural studies of proteins containing the classical Cys_2His_2 zinc finger motif have led to novel insights into mechanisms of DNA binding and to a better understanding of their broader functions in transcriptional regulation.

Addresses
Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA
*e-mail: wright@scripps.edu

Current Opinion in Structural Biology 2001, 11:39–46

0959-440X/01/$ – see front matter

Abbreviations

Arf	ADP-ribosylation factor
ATF-2	activating transcription factor 2
BIR	baculoviral IAP repeat
CR	cysteine-rich
CRE	cyclic AMP response element
FOG	friend of GATA
GAP	GTPase-activating protein
IAP	inhibitor of apoptosis
MRE	metal-response element
MTF-1	MRE-binding transcription factor 1
NC	nucleocapsid
PDB	Protein Data Bank
PI(3)P	phosphatidylinositol 3-phosphate
TFIIS	transcription factor IIS
TFIIIA	transcription factor IIIA
WT1	Wilms' tumor suppressor protein

Introduction

The zinc finger was first recognized 15 years ago as a repeated zinc-binding motif, containing conserved cysteine and histidine ligands, in *Xenopus* transcription factor IIIA (TFIIIA) [1]. Since that time, numerous other zinc-binding motifs have been identified and designated as zinc fingers. These vary widely in structure, as well as in function, which ranges from DNA or RNA binding to protein–protein interactions and membrane association. We now recognize the classical Cys_2His_2 zinc finger as the first member of a rapidly expanding family of zinc-binding modules. For the purposes of this review, we define a zinc finger to be any small, functional, independently folded domain that requires coordination of one or more zinc ions to stabilize its structure. In this article, we review recent advances in the structural biology of zinc finger proteins. We focus first on zinc fingers that function by binding nucleic acids or otherwise participate in transcriptional or translational processes. However, for the sake of completeness, we extend the discussion to encompass other novel zinc finger domains whose structures have been elucidated recently, even though their functions do not involve interactions with nucleic acids. A notable omission from this review is the zinc fingers from nuclear hormone receptors, which are reviewed in another article in this issue (see the review by Rastinejad, pp 33–38).

DNA binding by Cys_2His_2 zinc fingers

Proteins containing the classical Cys_2His_2 zinc finger are amongst the most abundant in eukaryotic genomes. Many of these proteins are transcription factors that function by recognition of specific DNA sequences. Since the structure of a single zinc finger was first reported in 1989 [2] (Figure 1a), the structures of numerous complexes with DNA have been described, beginning with that of Zif268 [3]. These structures have established the invariance of the $\beta\beta\alpha$ framework of the Cys_2His_2 zinc finger module and have provided a sound basis for understanding the nature of the interactions that mediate DNA binding. Specificity is determined by sidechain–base interactions involving a discrete set of residues located at the fingertip or on the surface of the helix. Additional interactions with the phosphate backbone and between adjacent zinc fingers contribute to binding affinity. The DNA-binding functions of zinc finger proteins have recently been the subject of an extensive review [4••].

Approximately half of the known Cys_2His_2 zinc finger proteins contain a highly conserved linker of sequence TGEKP that connects adjacent fingers [4••,5••]. A long-standing question about zinc finger proteins has been whether the linkers play an active or passive role in DNA binding. Early mutagenesis studies suggested that the linkers do play a role in DNA binding, with single site mutations reducing the binding affinity by as much as 20-fold [6]. A recent comparison of the structures of TGEKP linkers in the available X-ray and NMR structures of zinc finger proteins in complex with DNA revealed that the linker conformation is remarkably conserved [5••]. Although it has long been known that one of the functions of the TGEKP linker is to cap the C terminus of the preceding helix [3,7], it has only recently been recognized that this C-capping interaction is not present in the free protein and is induced upon binding of the zinc finger protein to DNA [5••,8]. Dynamic NMR measurements on zinc finger domains of TFIIIA and of the Wilms' tumor suppressor

protein (WT1) [9,10••] revealed that the linker is dynamically disordered in the free protein, but adopts a well-defined structure with restricted backbone flexibility upon binding to DNA.

A survey of zinc finger protein sequences in the SWISS-PROT database shows that more than 70% of the linkers can potentially form a C-capping motif, suggesting that DNA-induced helix capping plays an important and general role in zinc finger protein function [5••]. A model for DNA binding has been proposed in which it is assumed that the intrinsic flexibility of the linkers will facilitate diffusion along the DNA as a zinc finger protein searches for its cognate binding site. Once the correct DNA sequence is encountered, structure is induced in the linkers, which then act as inducible 'snaplocks', holding the fingers in the proper orientation for optimal interactions in the major groove [5••]. Recognition of the important functional role played by conserved capping motifs in linker sequences may facilitate the identification of subsets of DNA-binding fingers from the genome sequences.

Modification of linker sequences *in vivo* can have profound functional and physiological consequences. An excellent example of this is WT1, in which alternate splicing results in the insertion (+KTS) or omission (–KTS) of the tripeptide Lys-Thr-Ser in the canonical TGEKP linker between the third and fourth zinc fingers. The relative abundance of these isoforms is tightly regulated in the cell and variations are associated with human disease: for example, Frasier syndrome is caused by an intronic mutation that prevents production of the +KTS isoform [11]. The presence or absence of the KTS insert modulates both the DNA-binding affinity and the functional distribution of WT1 within the nucleus. The –KTS isoform binds DNA with high affinity and regulates transcription; in contrast, the +KTS variant binds DNA only weakly and associates preferentially with the splicing machinery, where it may interact with RNA [12]. A recent NMR study of the +KTS and –KTS isoforms bound to DNA has revealed the molecular basis by which alternate splicing mediates the biological function of WT1 [10••]. The KTS insertion increases the flexibility of the linker between fingers three and four, and abrogates binding of the fourth zinc finger to its cognate site in the DNA major groove, significantly lowering the DNA-binding affinity. Thus, modification of linker sequences by alternate splicing represents a novel mechanism whereby a single zinc finger gene can be utilized to fulfill different functions in the cell.

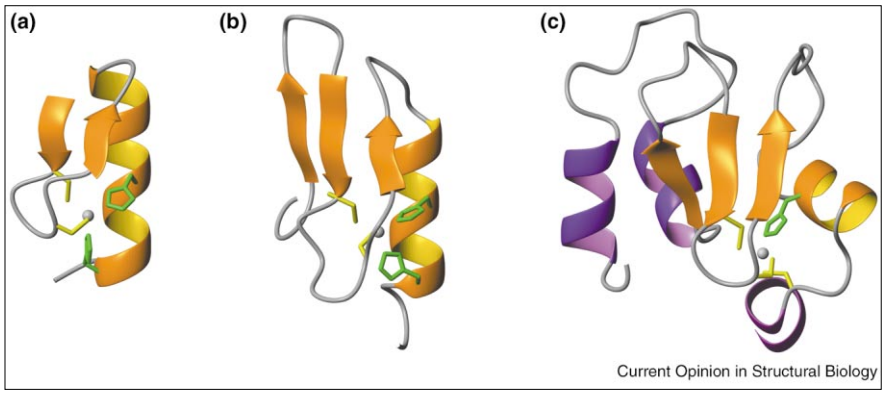
Zinc finger engineering

In recent years, much effort has been focused on the design of novel Cys_2His_2 zinc finger proteins that can specifically target unique binding sites within the human genome. Such 'designer' zinc fingers have important applications as probes and may ultimately prove valuable for human gene therapy. As a thorough review of recent progress in zinc finger engineering has been presented recently [13], this topic will be discussed only briefly. One of the current challenges in zinc finger protein engineering is in finding methods to link subsets of zinc fingers together to achieve the high binding affinity and exquisite specificity that would be required for applications in gene therapy. The binding affinity of Cys_2His_2 zinc finger proteins increases approximately 1000-fold for each additional finger connected by canonical TGEKP linkers. However, only modest improvements in affinity occur when proteins with more than three fingers are connected by TGEKP linkers; much effort has therefore been focused on the identification of suitable linkers for connecting three-finger modules (reviewed by Choo and Isalan [13], and Wolfe *et al.* [4••]). An alternative approach is to introduce dimerization motifs to facilitate the assembly of multiple zinc fingers on DNA. The utility of this approach has been demonstrated very recently by the construction of a fusion protein containing the GCN4 leucine zipper dimerization motif and fingers two and three from Zif268 [14••]. This four-finger chimera binds DNA with similar affinity and specificity as the best six-finger construct designed to date. The limitation of this chimeric protein is the relatively small size of the DNA site recognized (10 base pairs), but additional fingers can presumably be incorporated to produce a protein that recognizes longer DNA sequences.

Zinc-sensing proteins

Regulatory mechanisms have evolved in all organisms to control the homeostasis of metal ions. The response of mammalian cells to surplus zinc appears to be largely mediated by metallothionein, which acts to chelate and sequester the excess metal from the cellular environment. Metallothionein synthesis is induced by zinc excess and is regulated through the metal-response element (MRE) by the transcription factor MTF-1 (MRE-binding transcription factor 1), which contains six Cys_2His_2 zinc fingers. In yeast, the cellular response to zinc deficiency is regulated by the metal-sensing transcriptional activator Zap1, which contains seven Cys_2His_2 zinc fingers. Recent papers have begun to provide insights into the structural basis of zinc sensing by MTF-1 [15,16•] and Zap1 [17]. It appears that both proteins contain a core DNA-binding domain comprising a subset of the zinc fingers. In Zap1, the DNA-binding and zinc-responsive domains overlap and map to a set of five zinc fingers at the C-terminal end of the protein [17]. In contrast, fingers two to four of MTF-1 are involved in DNA binding, whereas finger one appears to function as a unique metal-sensing domain that prevents MTF-1 from binding to the MRE in the absence of exogenous zinc [16•]. Finger one deletion mutants bind DNA constitutively, but do not respond to zinc induction. The zinc-dependent molecular interactions involving finger one of MTF-1 that control DNA binding to the MRE have not been elucidated. Fingers five and six of MTF-1, which are highly conserved but are not part of the core DNA-binding region, appear to bind zinc with a relatively low affinity and to lack the typical $\beta\beta\alpha$ structural motif characteristic of Cys_2 His_2 zinc fingers [15]. However, it is

Figure 1



Variations on the classical Cys_2His_2 zinc finger. (a) Classical TFIIIA-type zinc finger (PDB code 1ZNF), (b) N-terminal zinc finger of SWI5 (PDB code 1NCS) and (c) BIR2 domain of XIAP (PDB code 1C9Q). The β sheet and α helix structural motif common to each structure is highlighted in orange. Additional helices in BIR2 are shown in magenta. All figures were generated from PDB files using MOLMOL for rendering [55].

not yet known whether these fingers play a metalloregulatory role. Indeed, much work is still needed to elucidate the structure of these proteins and to determine the mechanism by which they regulate transcription in response to changes in metal ion concentration. What is already clear, however, is that early suggestions [18] that zinc finger proteins might regulate gene expression in response to fluctuations in zinc concentration were prescient.

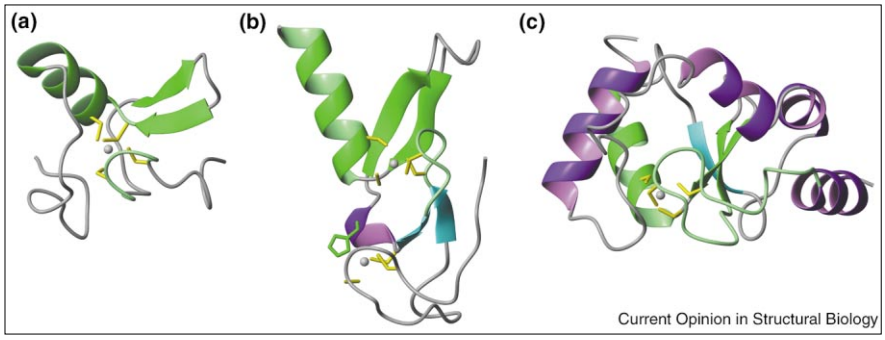
A broader function for TFIIIA-type ββα folds

Although the majority of the Cys_2His_2 zinc finger proteins identified to date are implicated in nucleic acid binding, it is becoming increasingly clear that some members of this superfamily function by mediating protein–protein interactions. For example, the zinc finger protein Ikaros, which plays a crucial role in lymphoid differentiation, forms homodimers through the association of the two C-terminal Cys_2His_2 zinc finger motifs [19]. The Ikaros-related protein Aiolos both homodimerizes and forms heterodimers with Ikaros through a two-zinc finger domain [20].

Zinc fingers also mediate functional interactions between the GATA-1 transcriptional activator (Figure 2a), which contains two CCCC zinc fingers, and the friend of GATA (FOG) protein. The FOG family of proteins includes the erythroid FOG-1 and the widely expressed FOG-2, as well as the *Drosophila* U-shaped protein, all of which contain eight or nine zinc fingers homologous to the classical TFIIIA-type Cys_2His_2 zinc fingers. In five of the zinc fingers in each protein, however, the last histidine ligand is replaced by cysteine [21,22•]. The interactions between these proteins involve the N-terminal zinc finger of GATA-1 and three to four of the CCHC zinc fingers of FOG [22•]. The Cys_2His_2 zinc fingers of FOG play no part in binding GATA-1 and are presently of unknown function. The residues involved in DNA binding and FOG binding are located on different faces of the N-terminal zinc finger of GATA-1, suggesting that it might simultaneously participate in recognition of GATA promoter sites and binding to the FOG cofactors [23]. The structure of a CCHC zinc finger from U-shaped has been reported and is very similar to that of the classical Cys_2His_2 motif [22•] (see also Update). Thus, substitution of a histidine ligand by cysteine appears to have little effect upon the structure.

The FOG CCHC zinc fingers are the first example of naturally occurring proteins containing functional substitutions of ligands within the Cys_2His_2 motif. The structure of this motif seems quite robust and ligand substitutions can apparently be accommodated without gross structural distortions. Indeed, it has recently been reported that the replacement of all of the cysteine ligands by histi-

Figure 2



The structures of the FYVE and Arf-GAP domains all contain a zinc-binding motif similar in structure to that from GATA-1 [56]. (a) GATA-type zinc finger from GATA-1 (PDB code 1GAT), (b) FYVE domain of Vps27p (PDB code 1VFY) and (c) Arf-GAP domain (PDB code 1DCQ). The GATA homology region is shown in green, with additional helices and β strands colored magenta and cyan, respectively.

dine in the three zinc fingers of Sp1 gives rise to a protein that is fully functional in DNA binding [24]. Structural studies of one of the mutant fingers show that the protein can accommodate the four histidine ligands with only minor distortions of the classical zinc finger fold.

An intriguing protein-interacting Cys_2His_2 finger is found in the transactivation domain of activating transcription factor 2 (ATF-2) or cyclic AMP response element (CRE) protein-1 (CRE-BP1) as it is also known. ATF-2 is a bZip transcription factor that binds to the CRE in response to increased levels of cAMP, both as a homodimer and as a heterodimer with c-Jun. The structure of the ATF-2 transactivation domain has been determined recently and reveals a folded N-terminal region that contains a Cys_2His_2 zinc finger coupled to an unstructured C-terminal subdomain [25]. The zinc finger domain adopts the characteristic $\beta\beta\alpha$ fold. The residues that constitute the hydrophobic core are highly conserved, whereas the surface charge distribution is very different from that of zinc fingers that are known to function through interactions with nucleic acids.

Nucleic acid recognition by other zinc finger motifs

The DM motif is a zinc-dependent DNA-binding domain found in a number of transcription factors that regulate sexual differentiation. A recent NMR structure revealed that the DM motif is a novel zinc finger containing intertwined CCHC and HCCC ligands [26•]. The two zinc-binding sites stabilize a compact globular domain that contains two α helices. The second helix extends into an unstructured C-terminal tail that functions as a nascent DNA-recognition helix, folding into a helical structure upon binding to DNA. The DM domain binds cooperatively to DNA as a dimer, making contacts in the minor groove. Mutations in both the zinc-binding motif and the C-terminal tail are associated with intersex disorders [26•]. The DM motif thus represents a novel zinc finger that, unlike most other zinc finger domains, recognizes the DNA minor groove.

The retroviral nucleocapsid (NC) protein from HIV-1 plays a major role in the recognition and packaging of the retroviral genome. The NC protein contains two zinc 'knuckles' separated by a seven-residue linker that, together with a long N-terminal tail, are unstructured in the free protein [27]. The protein recognizes four different stem loop structures within the HIV-1 packaging sequence, the Ψ-site. Structures of the NC protein bound to both the SL3 and SL2 RNA stem loops from the Ψ-site have now been reported [28••,29], providing important insights into the mechanism of RNA recognition. Both SL2 and SL3 contain disordered G-rich tetraloops, GG(U/A)G, which fold upon binding to the zinc knuckles. Upon binding to SL3, the N-terminal tail of the NC protein folds into a 3_{10} helix that binds in the major groove and makes contacts with the phosphodiester backbone [29]. The interactions with SL2 are quite different, however. SL2 contains a novel A-U-A base triple, which opens up the minor groove [30]. In the complex with the NC protein, the N-terminal zinc knuckle binds in the minor groove and packs against the base triple [28••]. The N-terminal 3_{10} helix contacts the phosphodiester backbone, but does not bind in the major groove or make specific hydrogen bonding contacts, as it does in the SL3 complex. Thus, the NC protein binds to the different RNA stem loop structures by an adaptive mechanism, adopting a different conformation in each complex and making different intramolecular and intermolecular interactions. Adaptive binding is undoubtedly facilitated by the intrinsic flexibility in the linker and in the N-terminal tail.

New zinc finger folds

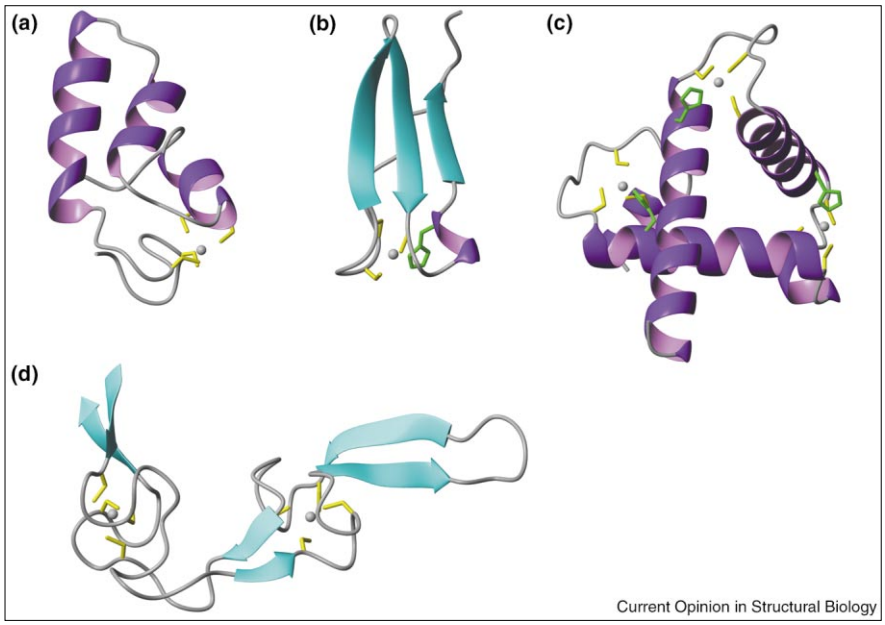
Zinc fingers in RNA polymerase

The crystal structure of yeast RNA polymerase II reveals six zinc-binding proteins, several of which appear to meet the criteria for classification as zinc fingers [31]. Subunit nine (Rpb9) forms two distinct zinc-binding domains separated by a long linker: the smaller C-terminal zinc-binding module has previously been shown to fold independently into a zinc ribbon motif similar to that observed in transcription factor IIS (TFIIS) [32,33]. Subunit Rpb12 appears to have a similar zinc ribbon structure in the intact polymerase [31] and displays sequence similarities with the zinc ribbon domain of Rpb9 [32]. Subunit Rpb10 binds zinc in an unusual CX_2CX_nCC motif, with two adjacent cysteine ligands. The structure of the isolated subunit has been determined by NMR [34••]. Folding is zinc dependent and the structure consists of a cluster of three helices (designated a zinc-bundle motif by the authors) with the zinc bound at one end (Figure 3a). Although these zinc finger subunits are essential for polymerase activity, it is not yet known whether they interact directly with nucleic acid or play some other role in the polymerase complex.

Ribosomal proteins

The 2.4 Å crystal structure of the large ribosomal subunit from *Haloarcula marismortui* contains four zinc-binding proteins homologous to the eukaryotic L24, L37, L37a and L44 proteins [35]. The structure of L24 is reminiscent of that of the GATA-type zinc finger. L37 and L37a bind zinc through a $CX_2CX_{11,14}CX_2C$ sequence motif and appear similar in structure to the zinc ribbon motif. The structure of the zinc-binding site in L44 also appears similar to the zinc ribbon, although there are large inserts in the zinc-binding motif, which has the sequence pattern $CX_5HX_{53}CX_2C$. The structure of ribosomal protein L36 was recently solved by NMR [36]. L36 also adopts the zinc ribbon fold, despite the substitution of the last cysteine ligand of the canonical zinc ribbon motif by histidine ($CX_2CX_{12}CX_4H$ coordination motif) (Figure 3b). The high propensity for zinc ribbon structures in the ribosome is intriguing and, as detailed analyses of the ribosome structure are completed, it will be of great interest to elucidate the nature of their interactions with RNA.

Figure 3



Ribbon diagrams of recently solved zinc finger structures. (a) Zinc bundle of Rpb10 (PDB code 1EF4), (b) zinc ribbon of L36 (PDB code 1DFE), (c) TAZ2 domain of CBP, and (d) CR domain of DnaJ (PDB code 1EXK). The coordinates of TAZ2 were obtained from the authors [38••]. α Helices and β strands are shown in magenta and cyan, respectively. Cysteine and histidine ligands are shown in yellow and green, respectively.

Transcriptional adaptor zinc fingers

The transcriptional adaptor proteins CBP and p300 contain two copies of a zinc finger motif, termed the TAZ finger [37], that are implicated in functional interactions with numerous transcription factors and viral oncoproteins. The NMR structure of the TAZ2 domain from CBP has been reported recently [38••]. The TAZ2 zinc finger folds into an unusual bundle of four helices that is stabilized by three zinc ions, each of which is bound to one histidine and three cysteine ligands in an HCCC motif (Figure 3c). Each zinc-binding site is formed from residues at the C terminus of an α helix, at the N terminus of the next helix and in the intervening loop. A peptide derived from the p53 activation domain binds to a specific site on the surface of the TAZ2 domain. The close sequence similarity between the TAZ2 and TAZ1 domains suggests that both are likely to adopt similar three-dimensional structures. Additional work will be required to elucidate the molecular features responsible for the differing ligand specificities of the two motifs.

Cysteine-rich domain of the chaperone DnaJ

The NMR structure of the cysteine-rich (CR) domain of *Escherichia coli* DnaJ reveals a novel zinc-binding motif [39•]. The protein binds two zinc ions using four CXXCXGXG peptide sequences (where X is any amino acid) that form an extended V-shaped structure (Figure 3d). The zinc coordination is interleaved, with each zinc ion coordinated by pairs of cysteine residues that are far apart in the sequence. As the spacing between the cysteine residues in the CR domain closely matches that of the Cys_4-type zinc fingers, it had been suggested that the structure may resemble that of the glucocorticoid receptor [40]. This is clearly not the case and the CR domain adopts a unique structure unlike that of any known zinc finger. The CXXCXGXG sequence motif is highly conserved from archaea to man, and is a topological signature for the CR domain. The domain appears to function in the binding of denatured proteins by DnaJ.

The BIR repeat

The baculoviral IAP (inhibitor of apoptosis) repeats (BIRs) contain a CCHC coordination motif. Many BIR domain proteins are involved in the regulation of programmed cell death, which they accomplish through inhibition of caspases. Structures have recently been determined for BIR domains from XIAP and c-IAP1 [41,42••,43], and from the anti-apoptotic protein survivin [44–46]. The conserved core of the BIR domain is made up of a central three-stranded β sheet and four short α helices. (Note that in the structure of the c-IAP1 BIR domain, the β sheet and one of the helices were not observed and the structure aligns poorly with the other five structures [41].) The zinc-binding motif in the BIR domain resembles that of the classical Cys_2His_2 zinc fingers, with the first two cysteine ligands located on a β hairpin and with the histidine ligand on the adjacent α helix [42••] (Figure 1c). Indeed, the structure is especially reminiscent of that of the first zinc finger domain of SWI5, with its three-stranded β sheet [47] (Figure 1b). However, the histidine-containing helix is considerably shorter in the BIR domain, terminating several residues before the final cysteine ligand. Also, the loop between the last β strand and the α helix is longer than in the classical zinc finger structures and is disordered in solution [43]. It is striking that a similar zinc-binding framework occurs in the classical Cys_2His_2 zinc finger and in the BIR domain, although it is adorned by additional helices in the

latter that mediate functional interactions. Indeed, despite the high degree of conservation of the core structure in the IAP and survivin BIR domains, differences in the N- and C-terminal helical regions are observed that probably relate to their differing functions.

Lipid-binding zinc fingers

The FYVE domain targets cytoplasmic proteins to specific membranes by recognition of phosphatidylinositol 3-phosphate [PI(3)P] through a highly conserved sequence motif. The FYVE domain binds two zinc atoms, which are required for folding [48], with CCCH and CCCC coordination [49•]. The ligands are interleaved in the sequence, with the first and third pair of ligands bound to one zinc and the second and fourth pair of ligands bound to the second zinc ion (Figure 2b). A structural resemblance between the second zinc-binding site of the FYVE domain and the GATA-1 (Figure 2a) and protein kinase C zinc fingers has been noted [49•]. Dimerization is required for high-affinity binding to PI(3)P-containing membranes [48]. Antiparallel homodimers of the FYVE domain of Hrs have been observed in a recent crystal structure and have given rise to a plausible model for membrane binding [50]. The structure of the Rab3A effector domain of rabphilin-3A [51•] is very similar to that of the FYVE domain. Although it has been suggested that the rabphilin zinc finger domain may also interact with phospholipids [51•], it does not contain the PI(3)P-binding sequence motif typical of FYVE domains and direct lipid binding is yet to be demonstrated.

Interestingly, a GATA-type zinc-binding motif is present in the Arf-GAP (ADP-ribosylation factor GTPase-activating protein) domain of the centaurin β family of membrane-associated proteins. Arf family GTPase proteins regulate membrane trafficking and contain pleckstrin homology (PH) domains that function together with the Arf GAP domain to mediate binding to phosphoinositides [52]. The structure of the Arf GAP domain has been reported recently and reveals an exposed zinc-binding motif that is similar in structure to the GATA-1 and protein kinase C zinc fingers [53•,54] (Figure 2c). The role of the zinc finger motif in phospholipid binding remains to be established.

Conclusions

Recent structural studies of zinc finger proteins have shed new insights into their extraordinary diversity of structure and function. It is chastening to realize, however, that of the large number of putative zinc finger motifs that have been identified, only a handful have been characterized structurally. Although some of these will undoubtedly prove to have novel folds, it is notable that recently determined structures of several previously uncharacterized zinc finger domains show that they are built on common structural cores, first seen in DNA-binding zinc fingers (the Cys_2His_2 motif [Figure 1a], GATA-1 [Figure 2a] and the β-ribbon zinc finger motif of TFIIS [Figure 3b]). Despite the remarkable progress made, much remains to be done to elucidate the structures, functions and mechanisms of action of zinc finger proteins. What is already very clear, however, is that these small, independently folded protein domains play a vital role in regulating a remarkable array of biological functions. Even the classical Cys_2His_2 zinc finger motif can no longer be viewed simply as a nucleic-acid-binding motif, but is now seen to perform additional functions beyond DNA and RNA recognition and packaging.

Update

The solution structure of the plant homeodomain (PHD) motif has been reported recently [57•,58]. It folds into an interleaved zinc finger domain that binds two zinc atoms. The zinc-binding core is similar in structure to that of the FYVE and RING domains. The PHD motif is found in both single and multiple copies within transcriptional control proteins. Mutations within the PHD motif are associated with developmental disorders, such as Williams–Beuren syndrome, X-linked mental retardation and autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy.

Further details of the FOG zinc finger structure have been recently described, along with chemical shift mapping and mutational analysis of the interaction with the N-terminal GATA-1 zinc finger [59]. The FOG zinc finger resembles the classical TFIIIA zinc finger, except for the substitution of the last zinc ligand histidine for cysteine.

In recent work on zinc finger engineering, a zinc finger protein that is based on the DNA-binding domain of Zif268 was optimized using phage display techniques to selectively recognize a human telomeric DNA sequence that forms a four-stranded guanine quadruplex [60]. Although this engineered human telomere-specific protein has potential diagnostic and therapeutic applications, further structural characterization is necessary to identify the mode of binding and specific intermolecular interactions within the complex.

Acknowledgements

JHL and BML are supported by fellowships from the National Institutes of Health (NIH). Research on zinc finger proteins is supported by NIH grant GM36643 to PEW.

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. Miller J, McLachlan AD, Klug A: Repetitive zinc-binding domains in the protein transcription factor IIIA from ***Xenopus*** oocytes. *EMBO J* 1985, 4:1609-1614.

2. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE: Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* 1989, 245:635-637.

3. Pavletich NP, Pabo CO: Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* 1991, 252:809-817.

4. Wolfe SA, Nekludova L, Pabo CO: DNA recognition by Cys_2His_2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* 2000, 29:183-212.
•• A thorough review of classical zinc finger proteins. Topics covered include DNA recognition and binding properties, as well as structure-based design developments.

5. Laity JH, Dyson HJ, Wright PE: DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. *J Mol Biol* 2000, 295:719-727.
•• This paper provides a structural explanation for the high sequence conservation of the consensus TGEKP linker in Cys_2His_2 zinc finger proteins. NMR data provide evidence for helix capping interactions that are induced upon binding to DNA. A model mechanism for DNA binding is proposed in which intrinsic linker flexibility in the free state facilitates the search for the target sequence, after which the DNA-induced linker structure functions as a 'snaplock' to hold adjacent fingers in the correct orientation in the major groove.

6. Choo Y, Klug A: A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res* 1993, 21:3341-3346.

7. Wuttke DS, Foster MP, Case DA, Gottesfeld JM, Wright PE: 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* 1997, 273:183-206.

8. Laity JH, Chung J, Dyson HJ, Wright PE: Alternative splicing of Wilms' tumor suppressor protein modulates DNA binding activity through isoform-specific DNA-induced conformational changes. *Biochemistry* 2000, 39:5341-5348.

9. Foster MP, Wuttke DS, Radhakrishnan I, Case DA, Gottesfeld JM, Wright PE: Domain packing and dynamics in the DNA complex of the N-terminal zinc fingers of TFIIIA. *Nat Struct Biol* 1997, 4:605-608.

10. Laity JH, Dyson HJ, Wright PE: Molecular basis for modulation of biological function by alternate splicing of the Wilms tumor suppressor protein. *Proc Natl Acad Sci USA* 2000, 97:11932-11935.
•• The structural basis for the different biological functions of two isoforms of WT1 was determined from NMR relaxation measurements of free and DNA-bound forms of the protein. A three amino acid KTS insertion in the linker between the third and fourth Cys_2His_2 zinc fingers, which is completely conserved in all vertebrate forms of WT1, abrogates the DNA binding by the fourth finger. The resultant lower binding affinity of this +KTS isoform probably dictates its putative post-transcriptional regulatory role.

11. Barbaux S, Niaudet P, Gubler MC, Grunfeld JP, Jaubert F, Kuttenn F, Fekete CN, Souleyreau-Therville N, Thibaud E, Fellous M, McElreavey K: Donor splice-site mutations in WT1 are responsible for Frasier syndrome. *Nat Genet* 1997, 17:467-470.

12. Larsson SH, Charlieu JP, Miyagawa K, Engelkamp D, Rassoulzadegan M, Ross A, Cuzin F, van Heyningen V, Hastie ND: Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell* 1995, 81:391-401.

13. Choo Y, Isalan M: Advances in zinc finger engineering. *Curr Opin Struct Biol* 2000, 10:411-416.

14. Wolfe SA, Ramm EI, Pabo CO: Combining structure-based design with phage display to create new Cys(2)His(2) zinc finger dimers. *Structure* 2000, 8:739-750.
•• The first reported chimeric protein designed to bind a continuous 10–12 base pair DNA site as a dimer with high specificity and femtomolar affinity. The fusion protein was constructed from the dimerization domain of the leucine zipper protein GCN4 fused to a tandem of Zif268 Cys_2His_2 zinc fingers (fingers two and three). The leucine zipper motif provided a rigid dimer interface that joined monomeric GCN4–Zif268 domains.

15. Chen X, Chu M, Giedroc DP: MRE-Binding transcription factor-1: weak zinc-binding finger domains five and six modulate the structure, affinity, and specificity of the metal- response element complex. *Biochemistry* 1999, 38:12915-12925.

16. Bittel DC, Smirnova IV, Andrews GK: Functional heterogeneity in the zinc fingers of the metalloregulatory transcription factor, MTF-1. *J Biol Chem* 2000, 275:37194-37201.
• A study involving full-length MTF-1 and various missing finger mutants identified zinc fingers two to four as the 'core' DNA-binding domain and suggested that finger one is responsible for zinc sensing. Zinc-mediated transcription, controlled by finger one of MTF-1, was demonstrated *in vivo* in mouse, insect and yeast cells.

17. Bird AJ, Zhao H, Luo H, Jensen LT, Srinivasan C, Evans-Galea M, Winge DR, Eide DJ: A dual role for zinc fingers in both DNA binding and zinc sensing by the zap1 transcriptional activator. *EMBO J* 2000, 19:3704-3713.

18. Berg JM: Zinc finger domains: hypotheses and current knowledge. *Annu Rev Biophys Biophys Chem* 1990, 19:405-421.

19. Sun L, Liu A, Georgopoulos K: Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. *EMBO J* 1996, 15:5358-5369.

20. Morgan B, Sun L, Avitahl N, Andrikopoulos K, Ikeda T, Gonzales E, Wu P, Neben S, Georgopoulos K: Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *EMBO J* 1997, 16:2004-2013.

21. Tsang AP, Visvader JE, Turner CA, Fujiwara Y, Yu C, Weiss MJ, Crossley M, Orkin SH: FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 1997, 90:109-119.

22. Fox AH, Liew C, Holmes M, Kowalski K, Mackay J, Crossley M:
• Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *EMBO J* 1999, 18:2812-2822.
A classical zinc finger motif, with the last zinc ligand histidine replaced by cysteine, mediates protein–protein interactions with the N-terminal GATA-1 finger through conserved hydrophobic residues. This result shows the ability of the classical zinc finger motif to accommodate a ligand substitution and, consequently, to change specificity from DNA recognition to protein–protein binding. Moreover, the juxtaposition of CCHC zinc fingers and classical CCHH fingers within the multifingered FOG proteins suggests a dual role in protein–protein binding and DNA recognition by these and other transcriptional cofactors (see also Update).

23. Kowalski K, Czolij R, King GF, Crossley M, MacKay JP: The solution structure of the N-terminal zinc finger of GATA-1 reveals a specific binding face for the transcriptional co-factor FOG. *J Biomol NMR* 1999, 13:249-262.

24. Hori Y, Suzuki K, Okuno Y, Nagaoka M, Futaki S, Sugiura Y: Artificial zinc finger peptide containing a novel His4 domain. *J Am Chem Soc* 2000, 122:7648-7653.

25. Nagadoi A, Nakazawa K, Uda H, Okuno K, Maekawa T, Ishii S, Nishimura Y: Solution structure of the transactivation domain of ATF-2 comprising a zinc finger-like subdomain and a flexible subdomain. *J Mol Biol* 1999, 287:593-607.

26. Zhu L, Wilken J, Phillips NB, Narendra U, Chan G, Stratton SM,
• Kent SB, Weiss MA: Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers. *Genes Dev* 2000, 14:1750-1764.
The authors describe the solution structure of unique intertwined zinc fingers that recognize the DNA minor groove through concomitant folding of the C-terminal helix.

27. Lee BM, De Guzman RN, Turner BG, Tjandra N, Summers MF: Dynamical behavior of the HIV-1 nucleocapsid protein. *J Mol Biol* 1998, 279:633-649.

28. Amarasinghe GK, De Guzman RN, Turner RB, Chancellor KJ, Wu ZR,
•• Summers MF: NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition. *J Mol Biol* 2000, 301:491-511.
The solution structure of nucleocapsid protein bound to a second Ψ-site RNA hairpin demonstrates the flexibility of the tandem zinc knuckle to adapt to the variation in RNA structure found within the HIV genomic packaging signal.

29. 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.

30. Amarasinghe GK, De Guzman RN, Turner RB, Summers MF: NMR structure of stem-loop SL2 of the HIV-1 psi RNA packaging signal reveals a novel A-U-A base-triple platform. *J Mol Biol* 2000, 299:145-156.

31. Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD: Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 2000, 288:640-649.

32. Wang B, Jones DN, Kaine BP, Weiss MA: High-resolution structure of an archaeal zinc ribbon defines a general architectural motif in eukaryotic RNA polymerases. *Structure* 1998, 6:555-569.

33. Qian X, Jeon C, Yoon H, Agarwal K, Weiss MA: Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIS. *Nature* 1993, 365:277-279.

34. Mackereth CD, Arrowsmith CH, Edwards AM, McIntosh LP: Zinc
•• bundle structure of the essential RNA polymerase subunit RPB10 from ***Methanobacterium thermoautotrophicum***. *Proc Natl Acad Sci USA* 2000, 97:6316-6321.
This novel three-helix zinc bundle utilizes a unique motif of adjacent cysteines in the last helix for zinc coordination.

35. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* 2000, 289:905-920.

36. Hard T, Rak A, Allard P, Kloo L, Garber M: The solution structure of ribosomal protein L36 from ***Thermus thermophilus*** reveals a zinc-ribbon-like fold. *J Mol Biol* 2000, 296:169-180.

37. Ponting CP, Blake DJ, Davies KE, Kendrick-Jones J, Winder SJ: ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem Sci* 1996, 21:11-13.

38. De Guzman RN, Liu HY, Martinez-Yamout M, Dyson HJ, Wright PE:
•• Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CDP. *J Mol Biol* 2000, 303:243-253.
This paper reports the first structure of an intact transcriptional adaptor zinc finger domain and identifies the binding surface for the p53 activation domain. The TAZ domain has a novel helical fold stabilized by three zinc ions and provides a template for binding of numerous cellular transcription factors and viral oncoproteins.

39. Martinez-Yamout M, Legge GB, Zhang O, Wright PE, Dyson HJ:
• Solution structure of the cysteine-rich domain of the ***Escherichia coli*** chaperone protein DnaJ. *J Mol Biol* 2000, 300:805-818.
A new zinc-binding structural motif is reported, in which the protein binds two zinc ions through interleaved CXXCXGXG motifs to form an extended V-shaped structure.

40. Szabo A, Korszun R, Hartl FU, Flanagan J: A zinc finger-like domain of the molecular chaperone DnaJ is involved in binding to denatured protein substrates. *EMBO J* 1996, 15:408-417.

41. Hinds MG, Norton RS, Vaux DL, Day CL: Solution structure of a baculoviral inhibitor of apoptosis (IAP) repeat. *Nat Struct Biol* 1999, 6:648-651.

42. Sun C, Cai M, Gunasekera AH, Meadows RP, Wang H, Chen J,
•• Zhang H, Wu W, Xu N, Ng SC, Fesik SW: NMR structure and mutagenesis of the inhibitor-of-apoptosis protein XIAP. *Nature* 1999, 401:818-822.
The first complete structure of the inhibitor of apoptosis BIR domain shows structural similarity to classical zinc fingers and has a flexible N-terminal region that is linked to caspase-3 inhibition.

43. Sun C, Cai M, Meadows RP, Xu N, Gunasekera AH, Herrmann J, Wu JC, Fesik SW: NMR structure and mutagenesis of the third Bir domain of the inhibitor of apoptosis protein XIAP. *J Biol Chem* 2000, 275:33777-33781.

44. Verdecia MA, Huang H, Dutil E, Kaiser DA, Hunter T, Noel JP: Structure of the human anti-apoptotic protein survivin reveals a dimeric arrangement. *Nat Struct Biol* 2000, 7:602-608.

45. Chantalat L, Skoufias DA, Kleman JP, Jung B, Dideberg O, Margolis RL: Crystal structure of human survivin reveals a bow tie-shaped dimer with two unusual alpha-helical extensions. *Mol Cell* 2000, 6:183-189.

46. Muchmore SW, Chen J, Jakob C, Zakula D, Matayoshi ED, Wu W, Zhang H, Li F, Ng SC, Altieri DC: Crystal structure and mutagenic analysis of the inhibitor-of-apoptosis protein survivin. *Mol Cell* 2000, 6:173-182.

47. Neuhaus D, Nakaseko Y, Schwabe JW, Klug A: Solution structures of two zinc-finger domains from SWI5 obtained using two-dimensional 1H nuclear magnetic resonance spectroscopy. A zinc-finger structure with a third strand of beta-sheet. *J Mol Biol* 1992, 228:637-651.

48. Kutateladze TG, Ogburn KD, Watson WT, de Beer T, Emr SD, Burd CG, Overduin M: Phosphatidylinositol 3-phosphate recognition by the FYVE domain. *Mol Cell* 1999, 3:805-811.

49. Misra S, Hurley JH: Crystal structure of a phosphatidylinositol 3
• phosphate-specific membrane-targeting motif, the FYVE domain of Vps27p. *Cell* 1999, 97:657-666.
The first crystal structure of the FYVE domain reveals an interleaved zinc ligand conformation similar to, but not superimposable on, the C_3HC_4 RING domain. The FYVE domains mediate signal transduction by phosphatidylinositol 3-phosphate [PI(3)P] through a conserved binding motif. Modeling of a PI(3)P molecule within the structure of the FYVE domain helps to explain the high specificity for PI(3)P versus other phosphorylated phosphoinositides.

50. Mao Y, Nickitenko A, Duan X, Lloyd TE, Wu MN, Bellen H, Quiocho FA: Crystal structure of the VHS and FYVE tandem domains of Hrs, a protein involved in membrane trafficking and signal transduction. *Cell* 2000, 100:447-456.

51. Ostermeier C, Brunger AT: Structural basis of Rab effector
• specificity: crystal structure of the small G protein Rab3A complexed with the effector domain of rabphilin-3A. *Cell* 1999, 96:363-374.
The crystal structure of the rabphilin-3A effector domain is homologous to that of the FYVE domain, but lacks the phosphoinositide-binding motif.

52. Kam JL, Miura K, Jackson TR, Gruschus J, Roller P, Stauffer S, Clark J, Aneja R, Randazzo PA: Phosphoinositide-dependent activation of the ADP-ribosylation factor GTPase-activating protein ASAP1. Evidence for the pleckstrin homology domain functioning as an allosteric site. *J Biol Chem* 2000, 275:9653-9663.

53. Goldberg J: Structural and functional analysis of the ARF1
• ARFGAP complex reveals a role for coatomer in GTP hydrolysis. *Cell* 1999, 96:893-902.
The structure of GTPase activation domain Arf GAP has a zinc finger core with structural similarity to GATA-type fingers.

54. Mandiyan V, Andreev J, Schlessinger J, Hubbard SR: Crystal structure of the ARF-GAP domain and ankyrin repeats of PYK2-associated protein beta. *EMBO J* 1999, 18:6890-6898.

55. Koradi R, Billeter M, Wuthrich K: MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 1996, 14:51-55, 29-32.

56. Omichinski JG, Clore GM, Schaad O, Felsenfeld G, Trainor C, Appella E, Stahl SJ, Gronenborn AM: NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* 1993, 261:438-446.

57 Pascual J, Martinez-Yamout M, Dyson HJ, Wright PE: Structure of the
• PHD zinc finger from human Williams–Beuren syndrome transcription factor. *J Mol Biol* 2000, 304:723-729.
The first reported structure of a PHD (plant homeodomain) zinc finger motif reveals an unexpected structural similarity with the FYVE and RING finger motifs in the core zinc-binding region. Mutations of ligands in one of the zinc-binding sites and of a conserved residue located in a putative molecular interaction site are associated with human diseases.

58 Capili AD, Schultz DC, Rauscher FJ III, Borden KLB: Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. *EMBO J* 2001, 20:in press.

59 Liew CK, Kowalski K, Fox AH, Newton A, Sharpe BK, Crossley M, MacKay JP: Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that mediate protein-protein interactions. *Structure* 2000, 8:1157-1166.

60 Isalan M, Patel SD, Balasubramanian S, Choo Y: Selection of zinc fingers that bind single-stranded telomeric DNA in the G-quadruplex conformation. *Biochemistry* 2001, in press.