The EMBO Journal vol.14 no.23 pp.5947–5956, 1995

Novel topology of a zinc-binding domain from a protein involved in regulating early *Xenopus* development

Katherine L.B.Borden[1,2], John M.Lally[3], Stephen R.Martin[4], Nicola J.O'Reilly[5], Laurence D.Etkin[6] and Paul S.Freemont[3,7]

[1]Laboratory of Molecular Structure and [4]Division of Physical Biochemistry, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, [5]Peptide Synthesis Laboratory and [3]Protein Structure Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK and [6]Department of Molecular Genetics, University of Texas M.D.Anderson Cancer Center, Houston, TX 77030, USA
[2]Present address: Protein Structure Laboratory, Imperial Cancer Research Fund, London WC2A 3PX, UK

[7]Corresponding author

***Xenopus* nuclear factor XNF7, a maternally expressed protein, functions in patterning of the embryo. XNF7 contains a number of defined protein domains implicated in the regulation of some developmental processes. Among these is a tripartite motif comprising a zinc-binding RING finger and B-box domain next to a predicted α-helical coiled-coil domain. Interestingly, this motif is found in a variety of proteins including several proto-oncoproteins. Here we describe the solution structure of the XNF7 B-box zinc-binding domain determined at physiological pH by ^{1}H NMR methods. The B-box structure represents the first three-dimensional structure of this new motif and comprises a monomer having two β-strands, two helical turns and three extended loop regions packed in a novel topology. The r.m.s. deviation for the best 18 structures is 1.15 Å for backbone atoms and 1.94 Å for all atoms. Structure calculations and biochemical data shows one zinc atom ligated in a Cys2–His2 tetrahedral arrangement. We have used mutant peptides to determine the metal ligation scheme which surprisingly shows that not all of the seven conserved cysteines/histidines in the B-box motif are involved in metal ligation. The B-box structure is not similar in tertiary fold to any other known zinc-binding motif.**

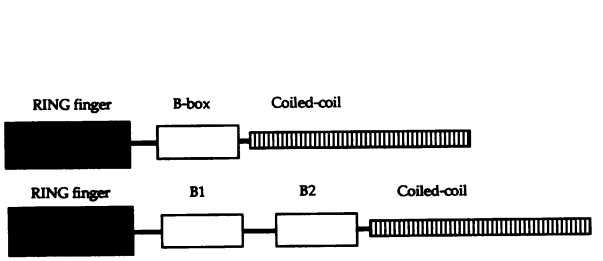
Keywords: B-box domain/NMR/protein structure/RING finger/*Xenopus* XNF7

Introduction

Xenopus nuclear factor, XNF7 (609 amino acids), is a maternally expressed protein that is retained in the cytoplasm until the mid-blastula transition, at which time it enters the nucleus (Dreyer *et al.*, 1983; Miller *et al.*, 1989, 1991; Reddy *et al.*, 1991; Li *et al.*, 1994a). Based on its primary structure, subcellular localization, nucleic acid binding properties and the ability of the N-terminal domain to transactivate a reporter gene, XNF7 is thought to function as a transcription factor (Miller *et al.*, 1989; Reddy *et al.*, 1991). Preliminary evidence indicates that XNF7 is involved in dorsal ventral patterning of the embryo (W.Shou and L.Etkin, unpublished observations). Interestingly, XNF7 has been detected in association with the mitotic spindle and condensed chromosomes during mitosis (Li *et al.*, 1994b), a process which is functionally important.

XNF7 belongs to a family of proteins which contain two novel zinc finger domains named RING (Freemont *et al.*, 1991; Freemont, 1993; Lovering *et al.*, 1993) and B-box (Reddy and Etkin, 1991) followed closely (5–8 amino acids) by a predicted α-helical coiled-coil domain forming a tripartite motif (Figure 1a; Kastner *et al.*, 1992; Reddy *et al.*, 1992). The spacing between the three elements of the motif is highly conserved among family members, suggesting that the relative positions of the domains is of functional importance.

There are a number of proto-oncoproteins within the B-box family, including the ret finger protein (RFP) (Takahashi *et al.*, 1988), PML (de Thé *et al.*, 1991; Goddard *et al.*, 1991; Kakizuki *et al.*, 1991; Kastner *et al.*, 1992) and TIF1 (T18; Miki *et al.*, 1991; Kastner *et al.*, 1992; Le Douarin *et al.*, 1995). RFP, PML and T18 (TIF1) are oncogenic in humans and mice when found as translocations that include the RING, B-box and α-helical coiled-coil domains recombined with other genes. PML is of particular importance owing to its direct association



(b)

```
                    1                  10                  20                  30                  40
XNF7 (219-260)      RPLEKC  SEH   DERLKLYCKDDGTLSCVICRDSLK H      ASHNFLPI
PwA33 (233-274)     KPKEKC  DEH   DERLKLFCKDDGTLACVICRDSKL H      SNHNFLPI
RPT-1 (91-132)      QKVNIC  AQH   GEKLRLFCRKDMMVICWLCERSQE H      RGHQTALI
SS-A/Ro (87-128)    TQGERC  AVH   GERLHLFCEKDGKALCWVCAQSRK H      RDHAMVPL
RFP   (101-142)     GEMGVC  EKH   REPLKLYCEEDQMPICVVCDRSRE H      RGHSVLPL
ATDC B2 (220-260)   FEARKC  PVH   GKTMELFCQTDQTCICYLCMFQE H      KNHSTVTV
EFP B2 (152-191)    LLRRKC  SQH   NRLREFFCPEHSECICHICLVE  H      KTCSPASL
TIF1 B2 (217-258)   QRPVFC  PFH   KKEQLKLYCETCDKLTCRDCQLLE H      KEHRYQFI
PML B2 (184-232)    TNNIFCSNPNH RTPTLTSIYCRGCSKPLCCSCALLDSSH    SELKCDISAE
ATDC B1 (168-216)   SEEVLC  DSCIGNKQKAVKSCLVCQASFCELHLKP   HLEGAAFRDHQLLEP
EFP B1 (102-148)    NAQVAC  DHC   LKEAAVKTCLVCMASFCQEHLQP   HFDSPAFQDHPLQPP
TIF1 B1 (157-204)   KSNQVC  TSC   EDNAEANGFCVECVEWLCKTCIRA   HQRVKFTKDHTVRQK
PML B1 (124-166)    DAQAVC  TRC   KESADFWCFECEQLLCAKCFEA   H   QWFLKHEARPL
                    $*      *     $ $$*  *  $$*  *         *     * $ $$
```

Fig. 1. The B-box family. (**a**) Schematic representation of the tripartite motif for XNF7. The RING and B-box are separated by 39 amino acids. The B-box and α-helical coiled-coil domains are adjacent. (**b**) Sequence alignment of the B-box family. The putative conserved metal ligands are shown in bold and are labelled with a *. Conserved hydrophobic residues are labelled with a $. Each sequence is referenced in the text and the top numbering refers to the XNF7 B-box peptide (residues 1–42). The actual protein numbering is given in parentheses.

Table I. ^{1}H NMR assignments of the B-box peptide

Residue	Chemical shift (p.p.m.)			
	NH	αH	β/β'H	Others
Arg1	–	4.35	1.94	1.78 (γ), 3.25(δ)
Pro2	–	4.51	2.22/2.35	1.92/2.00 (γ/γ'), 3.62/3.78 (δ/δ')
Leu3	–	5.00	1.95	1.54 (γ), 1.04 (δ)
Glu4	8.53	3.78	2.25/2.89	2.71(γ)
Lys5	9.14	3.86	1.82	1.51(γ), 1.71 (δ), 2.96 (ε)
Cys6	8.28	4.14	–	
Ser7	7.74	4.55	4.16/4.29	
Glu8	6.86	4.84	2.28/2.36	
His9	8.33	3.96	3.19/3.49	8.27(H2), 6.86 (H4)
Asp10	9.18	4.28	3.06/3.19	
Glu11	8.54	4.14	1.45/1.85	
Arg12	7.74	4.35	1.94	1.70 (γ), 3.25 (δ)
Leu13	7.56	4.13	1.81/1.97	1.44 (γ), 0.74/0.84 (δ/δ')
Lys14	7.74	4.56	1.42/1.66	1.11 (γ), 1.51 (δ), 2.85 (ε)
Leu15	8.63	5.33	1.35/1.42	1.27 (γ), 0.33/0.79 (δ/δ')
Tyr16	9.78	5.00	2.73/2.85	6.85 (3,5), 6.95 (2,6)
Cys17	9.58	5.18	3.25/3.81	
Lys18	7.07	3.48	1.55	0.78/0.99 (γ/γ'), 2.98 (ε)
Asp19	7.01	4.45	2.42/2.95	
Asp 20	8.13	4.77	2.71/2.90	
Gly21	8.34	3.78/4.00		
Thr22	7.31	4.56	4.04	1.21 (γ)
Leu23	8.49	4.34	1.68	0.91/0.96 (δ/δ')
Ser24				
Cys25	8.82	5.46	2.69/3.58	
Val26	9.29	3.58	2.12	0.92/0.98 (γ/γ')
Ile27	8.94	3.68	1.08	0.28 (γ),1.16 (γMe), 0.73 (δ)
Cys28	8.23	4.00	2.77/3.67	
Arg29	8.48	4.25	1.85	1.65 (γ), 3.22 (δ)
Asp30	8.20	4.88	3.36/3.58	
Ser31				
Leu32	7.78	4.09	1.30/1.75	1.61 (γ), 0.91/0.95 (δ/δ')
Lys33	–	3.90	1.45/1.65	1.01/1.06 (γ/γ'), 1.60 (δ), 2.89 (ε)
His34	7.40	4.02	3.05/3.46	8.24 (H2), 7.47 (H4)
Ala35	7.37	4.30	1.42	
Ser36				
His37	–	4.07	3.70	8.18 (H2), 7.49 (H4)
Asn38	8.85	5.06	2.71	7.14,7.82 (NH_2A)/7.02,7.88 (NH_2B)[a]
Phe39	8.92	5.45	2.73/2.95	6.85 (4), 7.45 (3,5), 7.18 (2,6)
Leu40	9.61	5.08	1.76	1.67 (γ), 0.99/1.02 (δ/δ')
Pro41	–	4.86	2.18/2.38	2.04 (γ), 3.75/3.88 (δ/δ')
Ile42				

Assignments were carried out using standard assignment methods as described in the text. Data were collected at pH 7.5 and 30°C.
[a]There were two cross-peaks in the ^1H_2O TOCSY which corresponded to the amide side-chain of Asn38. The presence of two cross-peaks indicated that these protons were undergoing slow exchange on the NMR time-scale.

with acute promyelocytic leukaemia as a fusion protein with retinoic acid receptor alpha (e.g. see Goddard *et al.*, 1991). TIF1 has been shown to interact with several nuclear receptors *in vivo* and is proposed to mediate the ligand-dependent transcriptional activation function of nuclear receptors (Le Douarin *et al.*, 1995). EFP, an oestrogen-responsive gene product, is another family member thought to represent an oestrogen-responsive transcription factor mediating phenotypic expression due to oestrogen action (Inoue *et al.*, 1993). PML, TIF1 and EFP all possess two B-box-like domains (B1 and B2) which appear to form a subgrouping of the B-box family as highlighted by the presence of an extra potential metal ligand (Figure 1b; Reddy *et al.*, 1992; Inoue *et al.*, 1993). The other members of the family include RPT-1 (Patarca *et al.*, 1988), SS-A/Ro-52 (Chan *et al.*, 1991) and PwA33 (Bellini *et al.*, 1993). RPT-1 (mouse) has been implicated in the down-regulation of the interleukin-2 receptor, acting as a putative transcription factor (Patarca *et al.*, 1988). SS-A/Ro-52 is a human auto-antigen and forms part of a ribonucleoprotein complex, although no direct interaction between SS-A/Ro-52 and RNA has been observed (Chan *et al.*, 1991; Slobbe *et al.*, 1992). PwA33 (*Pleurodeles*) may also function as a component of RNP complexes (Bellini *et al.*, 1993). To date, two gene products have been reported as possessing a B-box and coiled-coil domain but with no RING finger. These include a candidate gene for ataxia–telangiectasia group D (ATDC) which possesses two B-box domains and a coiled-coil domain similar to PML, TIF1 and EFP (Leonhardt *et al.*, 1994). The other is a gene found at the breast cancer locus, 1 A1.3B (Campbell *et al.*, 1994) which has one B-box and a coiled-coil domain, and appears tightly linked to the BRCA1 gene which interestingly possesses a RING finger but no B-box or coiled-coil domains (Brown *et al.*, 1994; Miki *et al.*, 1994).

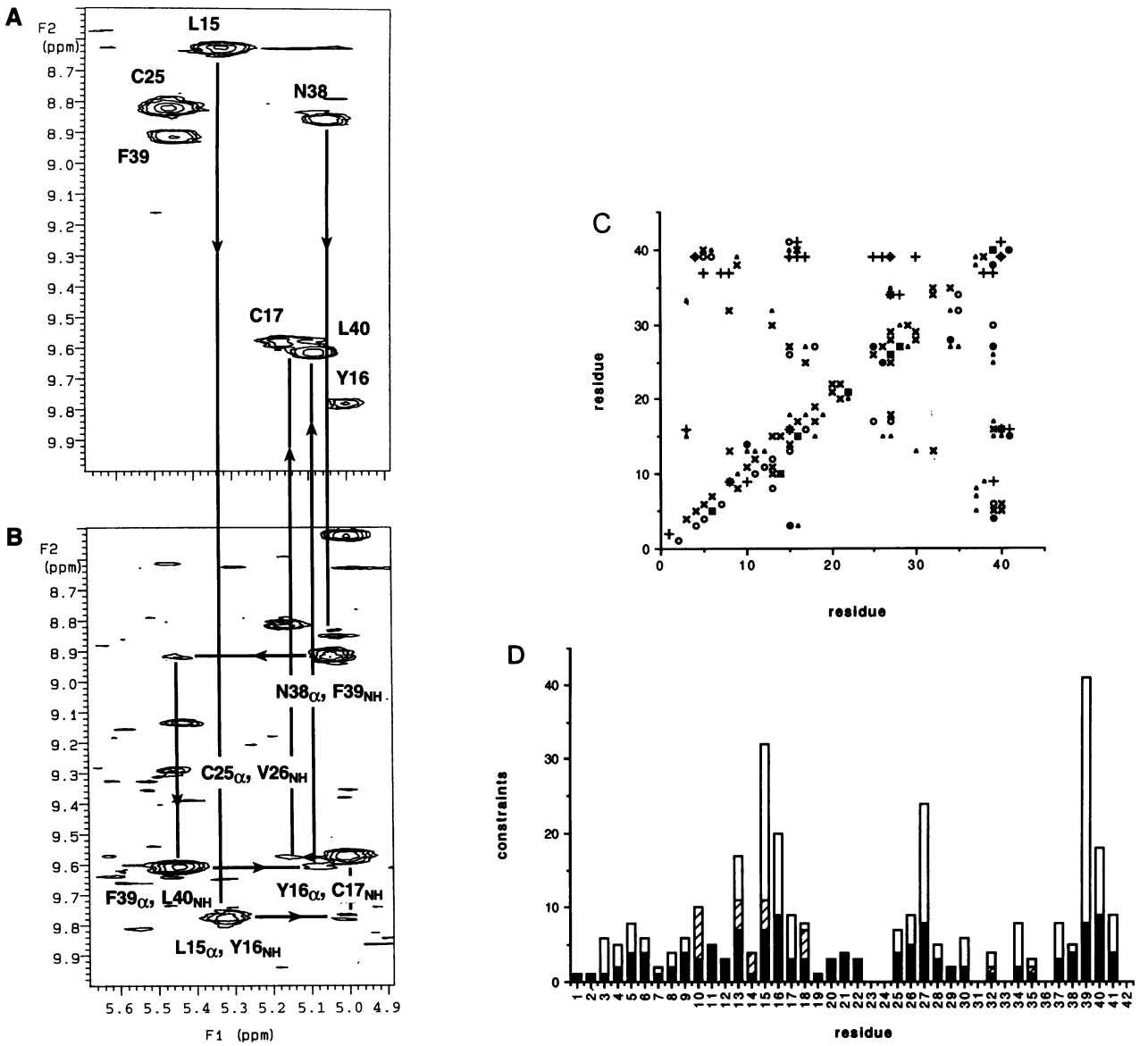


Fig. 2. NMR data used in XNF7 B-box structure calculations. (**A**) A section of the amide alpha region of a TOCSY and (**B**) a 290 ms NOESY in ^1H_2O. In (A) and (B) the $\alpha_i N_{i+1}$ connectivities observed in the β-strands are traced out and the corresponding positions of the intra-residue connectivities are indicated in the TOCSY spectrum. Most NOEs occur between the α-carbon proton from the preceding residue to the amide proton of the next residue. For example, L15αY16NH indicates a NOE between Leu15 α-carbon proton and Tyr16 amide proton. See text for further details. (**C**) A distance map summarizing the NOEs observed for the XNF7 B-box domain. Reading the residue from the *x*-axis indicates the proton type. Symbols: ○, α-carbon proton; ◆, side chain proton; ×, amide proton; +, ring proton. (**D**) Distance restraints per residue are shown with long-range (open bars), medium-range (cross-hatched bars) and short-range (solid bars).

In order to determine the molecular function of the B-box domain a number of studies have been initiated. Deletions of the XNF7 B-box domain result in a loss of binding to mitotic chromosomes (Li *et al.*, 1994b) while deletions and point mutations in the B-box of PwA33 results in the loss of association with the lampbrush loops of chromosomes in the oocyte nucleus in *Pleurodeles* (Bellini *et al.*, 1993; M.Bellini, personal communication). In addition, results of studies of the B-box family member RFP, which is closely related to XNF7 (Figure 1b), showed that the B-box and coiled-coil domains are both involved in homodimerization (T.Cao and L.Etkin, unpublished observations). Together, these data suggest that the B-box domain may mediate protein–protein interactions and in the case of XNF7 and PwA33 is important for the association of these proteins with subcellular structures during *Xenopus* and *Pleurodeles* development.

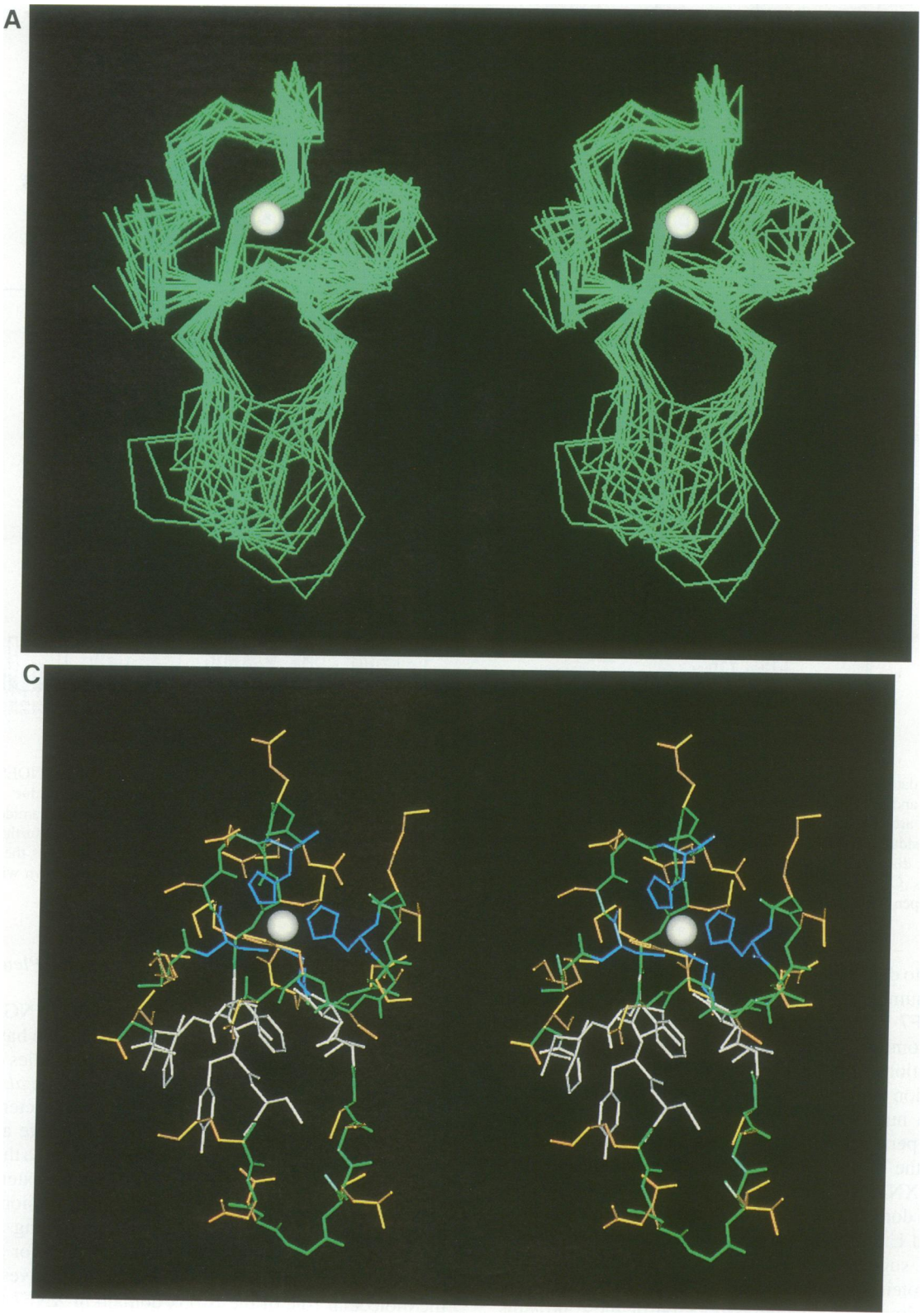
In order to understand the functions of the RING, B-box and coiled-coil domains in molecular detail, we have initiated a number of structural and biophysical studies of these domains (Borden *et al.*, 1993, 1995; Lovering *et al.*, 1993). Previously, we reported a biophysical characterization, including metal binding and secondary structure analysis, of a 42-residue synthetic peptide corresponding to the XNF7 B-box domain (Borden *et al.*, 1993). Here, we extend these earlier studies and now present the three-dimensional solution structure of the XNF7 B-box, representing the first structure, topology and metal ligation system for a B-box domain. The structure now provides a basis for investigating the molecular role of the B-box domain in XNF7 function

and *Xenopus* development and also provides insights into the function of the B-box domain within the B-box family.

Results and discussion

NMR structure determination

The three-dimensional structure of the XNF7 B-box domain (XNF7 residues 219–260) was determined using standard two-dimensional NMR techniques (see Materials and methods). It should be noted that experiments were carried out at physiological pH of 7.5 as the peptide—although soluble—was unstructured at lower pHs as assayed both by CD and NMR. ^{1}H chemical shift assignments are given in Table I. The NOESY footprint region (Figure 2B) shows connectivities from Leu15 to Cys17 and Asn38 to Leu40 which form part of the two β-strands in the B-box structure.



Also shown are the alignments to the corresponding region of the TOCSY spectra (Figure 2A). It should be noted that in the NOESY (Figure 2B) the amide proton (i) to α-carbon proton (i) intra-residue cross-peaks are missing whereas the α-carbon proton (i) to amide proton (i+1) inter-residue cross-peaks are very intense. This pattern of cross-peak intensity is characteristic of phi and psi angles which occur in extended conformation, such as β-strands (Wüthrich, 1986). In addition there are several tertiary NOEs between the strands (see Figure 2C). A summary of the observed sequential and long-range NOEs is given in Figure 2C in the distance map format. A total of 154 non-intra-residue distance constraints and 27 dihedral angle constraints (13 Phi, 7 Psi and 7 Chi1), were used as input into the program XPLOR (Brünger, 1992). Of the distance constraints, 64 were short-range (i to i+2), 12 medium (i+3 to i+5) and 78 long-range (i to > i+5). A summary of the constraints per residue is given in Figure 2D. It is notable that there are a number of areas within the peptide where there are no constraints. These regions generally correspond to positions where amide protons could not be assigned due to either overlap or exchange with bulk solvent. Of the 50 structures calculated, 39 converged with the remaining 11 as mirror images. The r.m.s. deviations for all 39 converged structures (residues 5–17; 25–41) are 1.39 Å for the backbone and 2.13 Å for all atoms. From the XPLOR calculations, 18 structures had NOE violations <0.5 Å with angle violations of <10° and have r.m.s. deviations (residues 5–17; 25–41) of 1.15 Å for the backbone and 1.94 Å for all atoms. An overlay of the α-carbon atoms of the 18 structures is shown in Figure 3A. The r.m.s. deviation for the flexible loop (residues 18–24) is 1.42 and 2.29 Å for backbone and all atoms respectively, suggesting that there is a correlated motion within the loop rather than complete disorder. The r.m.s. separation (residues 5–17; 25–41) between the average structure before and after inclusion of the Zn^{2+} parameters was 0.73 Å (backbone) and 1.13 Å (all atoms), demonstrating that the overall B-box structure did not change significantly after including Zn^{2+} in the structure calculations.

B-box structure and topology

The B-box structure consists of one β-strand (β1) and an extended region running parallel to β1 and a second β-strand (β2) running perpendicular to β1 (Figure 3B). The overall dimensions of the molecule are ~30×20×15 Å with the longest dimension including the flexible loop, described above, at its fullest extent (Figure 3B). A stereo view of all heavy atoms is given in Figure 3C.

The N-terminal three residues of the B-box peptide are poorly constrained and appear to be very flexible. Residues 4–6 form an extended region which continues into a loop/turn (residues 7–12) and then into the first β-strand region (β1; residues 13–18). β1 is followed by another extended loop and turn (residues 19–25) which appears to move in a correlated fashion (see Figure 3 and above). This loop/turn leads into an extended region which is parallel to β1. A helical turn (27–31) and extended region (residues 32–37) leads into the second β-strand (β2; residues 38–41) which lies perpendicular to β1. In the context of the tripartite motif, β2 leads directly into the predicted helical coiled-coil region (Figure 1A). This unusual topology may be necessary for the correct orientation of the B-box motif relative to the predicted coiled-coil domain.

The B-box domain has a tightly packed hydrophobic core consisting of the Leu15 side chain, Cys17 side chain, Ile27 side chain and Phe39 ring (Figure 3C). There appears to be a second smaller hydrophobic core formed by the Tyr16 ring packed on the Pro41 ring with involvement of the Leu40 side chain (Figure 3C). All of these hydrophobic interactions are probably necessary for stabilizing the overall B-box structure. There are also hydrogen bonds (as seen in the form of unexchanged amide protons)

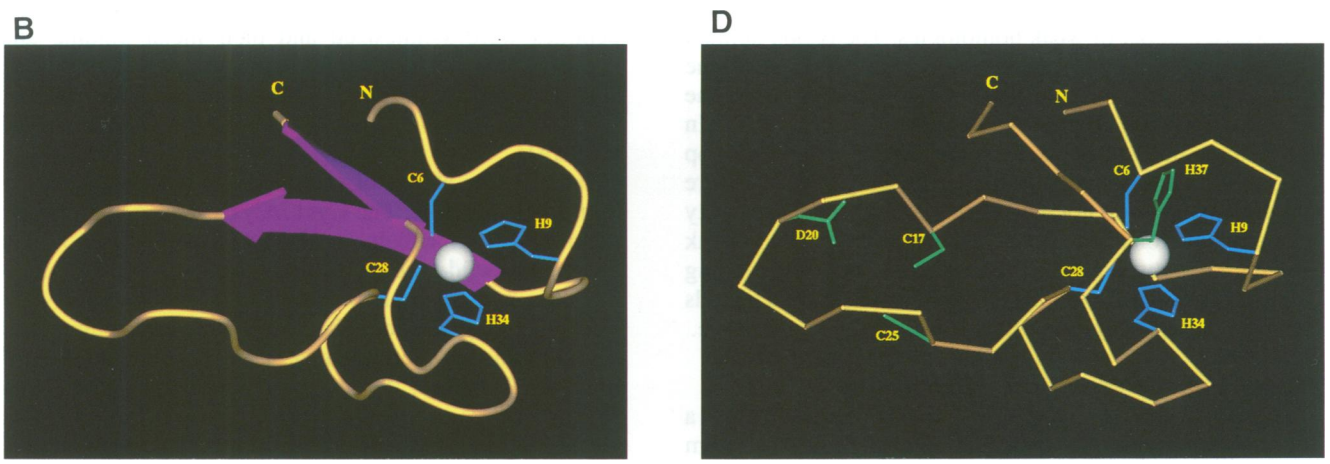


Fig. 3. The XNF7 B-box structure. (**A**) α-Carbon overlay of 18 XNF7 B-box structures. The three N-terminal residues are disordered and are not shown. The average position of the Zn^{2+} atoms are represented as a white sphere. The average NOE energy for the 18 structures was 66 ± 19 kcal/mol and the overall energy was 439 ± 61 kcal/mol. Deviations from idealized geometry were as follows: bonds 0.0165 ± 0.0005 Å, angles 1.17 ± 0.14° and impropers of 0.64 ± 0.12°. (**B**) Topology of the XNF7 B-box domain. The two β-strands are shown as arrows (magenta). The cysteine and histidine side chains which form the zinc binding site are labelled and shown in blue. The Zn^{2+} atom is represented as a white sphere and the N- and C- termini are labelled. (**C**) A stereo view showing all heavy atoms. Main chain atoms are shown in green, side chains in gold with core hydrophobic residues coloured white. The Zn^{2+} atom is represented as a white sphere with the proposed Zn^{2+} ligands coloured blue. (**D**) α-Carbon representation of the B-box structure in the same orientation as (B) showing the proposed metal ligands (Cys6, His9, Cys28 and His34; blue) and the other conserved putative ligands (Cys17, Cys25 and His37; green). The Zn^{2+} atom is represented as a white sphere and the N- and C- termini are labelled.

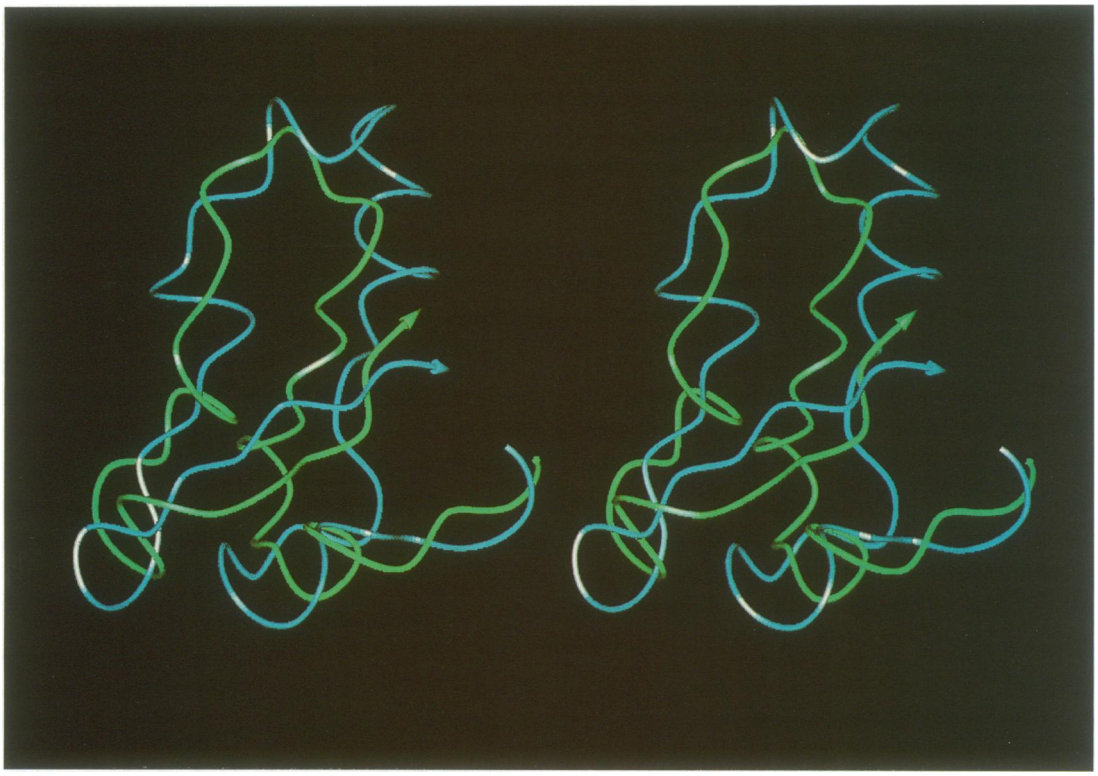


Fig. 4. Comparison of XNF7 B-box with nitrogenase iron binding protein. A stereo superposition of the B-box domain (residues 1–42) and nitrogenase iron binding protein (residues 40–88; r.m.s. separation for 26 C-α atoms is 2.6 Å). The main chain in both structures is represented as a thin tube with B-box in green and nitrogenase in blue.

between the carbonyl oxygen of Leu15 and the amide proton of Leu40, as well as the converse bond between the carbonyl of Leu40 and the amide proton of Leu15, which further stabilize the core of the B-box domain.

A search of representative structures and folds from the protein database for similarities to the B-box topology using the program STAMP (Russell and Barton, 1992) showed a number of weak homologies. The B-box domain has a similar chain tracing to a small portion of the nitrogenase iron binding protein but removed from the iron binding site (Figure 4; Georgiadis *et al.*, 1992). In particular, both proteins have a parallel extended loop region with a perpendicular region crossing over it (Figure 4). However, we do not believe that the observed similarity has any functional significance. Interestingly, there is weak homology between B-box and one of the iron binding sites in ferredoxin, although the zinc/iron metal and ligands do not overlap, suggesting little evolutionary significance.

Determination of the zinc binding site

Previous studies on the XNF7 B-box peptide using a chromophoric chelator indicated that only one zinc atom per peptide was bound (Borden *et al.*, 1993). Cobalt binding studies showed that the binding arrangement was tetrahedral and involved some cysteines. The curve shape in the 600–750 nm range also suggested a Cys2–His2 ligation scheme similar to that observed by Michael *et al.* (1992). However, which four of the seven possible ligands (see Figures 1B, 3B and D) were involved in metal ligation was unknown. Furthermore, a homologous B-box domain from the *Pleurodeles* protein PwA33, also binds one zinc atom per B-box molecule and uses His34 as a zinc ligand (M.Bellini, personal communication). Unfortunately, standard NMR techniques using ^{113}Cd could not be carried out as cadmium did not fully induce structure at concentrations used for circular dichroism (Borden *et al.*, 1993) and the peptide precipitated in the presence of ^{113}Cd at NMR concentrations. Therefore, Cys and/or His mutant peptides were synthesized and their metal binding and spectroscopic properties monitored. Eight peptides were studied: Cys17Cys25ΔAla; Asp20ΔAla; Cys28Asp20ΔAla; His34ΔAla; His37ΔAla; Cys6His9ΔAla; Cys6ΔAla and His9ΔAla.

Initially, the ability of each peptide to bind cobalt was monitored optically. The Cys17/Cys25 and Asp20 mutant peptides bound cobalt, producing a peak at ~650 nm similar to wild-type, which is characteristic of cobalt bound in a tetrahedral complex (Berg and Merkle, 1989). The remaining peptides, however, showed a diminished signal in this wavelength range, indicating less tetrahedral binding. Therefore, the ability of the mutant peptides to fold into native complexes was monitored by 1D ^{1}H NMR.

The 1D ^{1}H NMR spectra of each mutant peptide were collected in the absence and presence of zinc. In the wild-type spectra, four resonances are clearly visible in the region 5.1–6.0 p.p.m. (Borden *et al.*, 1993). These resonances correspond to the α-carbon protons of Leu40/Asn38, Cys17, Leu15 and Phe39/Cys25. The presence of α-carbon protons in this region of the spectra indicates the presence of β-strand (Wishart *et al.*, 1991). The Cys17/Cys25 and Asp20 mutant peptides had the same pattern and number of resonances in this region as

the wild-type spectrum with the exception that the Cys17 peak which was absent in the Cys17/Cys25 mutant, as one would expect. This suggests that the two peptides have native-like secondary structure. Further, the presence of the Phe39 α-carbon proton resonance is of particular importance when one considers that the Phe39 ring forms the core of the protein. The pattern of aromatic resonances of Cys17/Cys25 and Asp20 peptides is also similar to those observed in wild-type, suggesting that the tertiary structure of these peptides is similar to wild-type. In summary, Cys17, Cys25 and Asp20 appear not to be involved in metal ligation within the B-box monomer.

The remaining mutant peptides had fewer resonances in the 5.1–6.0 p.p.m. range. The Asp20/Cys28 mutation, only had one resonance present, corresponding to Leu15. The His34 and His37 mutant only had the Leu40/Asn38 resonance present. The Cys6/His9 mutant only had the Leu40/Asn38 and Leu15 resonances. This information clearly shows that none of these mutations is folding into a wild-type structure. Furthermore, the aromatic resonances for all these mutants were very similar to those observed for precipitated wild-type peptide which had been boiled in DTT in an attempt to recover it (data not shown). Therefore, it appears that mutations in any of these residues preclude the formation of a native-like structure.

Interestingly, His37 appears to be more flexible than the other putative ligands, suggesting that it may not be involved directly in ligation. The amide proton resonance is absent from both ^1H_2O TOCSY and NOESY experiments, suggesting that it is exposed to solvent and not involved in a stable hydrogen-bonding network. However, the amide protons are present in both ^1H_2O NOESY and TOCSY for all other putative ligands. Furthermore, the NH (γ) protons of the Asn38 side chain are involved in a slow exchange process (see Table I).

In summary, these studies indicate that there are five possible ligands for the tetrahedral binding site in XNF7 B-box, namely Cys6 (C6), His9 (H9), Cys28 (C28), His34 (H34) and His37 (H37). Of course, the assays used make it difficult to differentiate between mutations which destroy the peptides' ability to bind metal and therefore form a nascent structure or disrupt the tertiary contacts, particularly in the case of the histidine residues. From these results there are five possible ligation schemes: C6, H9, C28, H34; C6, C28, H34, H37; H9, C28, H34, H37; C6, H9, C28, H37 and C6, H9, H34, H37. Structural calculations were carried out in order to determine if the NOE and dihedral angle information obtained from the 2D ^1H NMR data of wild-type B-box could be used to ascertain the ligation scheme.

These data were used as input into the distance geometry embedding procedure with the appropriate metal ligation scheme included for each set of tests. These structures were used as input into simulated annealing protocols in XPLOR (Brünger, 1992). Twenty structures were generated for each ligation scheme. Table II summarizes the number of low violation structures out of the 20 calculated, and the average number of violations per residue and average overall energy for all 20 structures. The C6, H9, C28, H34 scheme produced a family of structures with low violations (<0.5 Å and <10°). The use of His34 as a zinc ligand is in agreement with zinc binding studies on PwA33 B-box homologue (M.Bellini, personal communication). Large violations were observed with all other ligation schemes with many fewer low violation structures and significantly higher energies than for the C6, H9, C28, H34 ligation scheme (Table II). Therefore 50 structures using the C6, H9, C28, H34 ligation scheme were calculated, the results of which are described above (see Figure 3). It is possible, however, that one of the alternative ligation schemes is correct, although it is clear that whichever of the schemes is used, there is little difference in the overall B-box fold since the r.m.s. deviation between representatives of each ligation scheme was 0.96 Å (backbone) and 1.03 Å (all atoms). The reasons for similarities between the various ligation schemes involving residues 6,9,28,34,37 is clear upon inspection of Figure 3D. Here, the four ligands shown in blue and His37 (in green) are located near to each other in space. Therefore only minor backbone rearrangements would be necessary to obtain one of the other ligation schemes. However, as shown in Table II, the structures calculated with His37 as a ligand only had 0–2 structures converged with <0.5 Å violations for distance restraints and <10° violations for angle constraints whereas the 6,9,28,34 ligation scheme yielded over four times the number of low violation structures. The data presented here indicate that the best fit of the 2D NMR data is obtained with the Cys6, His9, Cys28, His34 ligation scheme.

This leads to speculation as to what are the roles of the other putative metal-binding ligands (Figure 3D). One possible function may be metal-dependent inter-domain dimerization, especially with regard to Cys17 and Cys25 which occur on the same face and either side of the flexible loop and seem to form a half zinc site (Figure 3D). Interestingly, Asp20 (which is a cysteine in the B1/B2 sub-family) is also located on the same face of the loop (Figure 3D). In the context of the whole protein, the B-box domain leads directly into the predicted helical coiled-coil domain. Helical coiled-coil domains often form protein–protein interaction surfaces that can mediate either homo- or heterodimerization or even oligomerization (Harbury *et al.*, 1993). Preliminary studies with RFP show that both the B-box and coiled-coil regions together are necessary for efficient dimer formation (T.Cao and L.Etkin, unpublished observations). Therefore, in an XNF7 homodimer, two B-box domains would be positioned in close proximity to each other. One could speculate that the extra putative metal ligands orient the two domains into functionally active structural units, through inter-domain metal ligation as observed in the HIV Tat protein (Frankel *et al.*, 1988). His37 is located too far away from the loop with Cys17, Asp20 and Cys25 to interact directly or to share a ligand. The role of this residue may become more clear in the context of the entire protein. Clearly, the exact role of the remaining putative metal ligands will have to await further study.

Conserved residues within the B-box family

The main hydrophobic core residues 15, 16, 27, 28, 39, 40 and 41 are generally well conserved within the B-box family (Figures 1B and 3C). In particular, there is a striking preference for an aromatic residue preceding Cys17. In XNF7 this residue is a Tyr which is involved in interactions with Leu40 and Pro41, stabilizing the

Table II. Calculations of alternative ligation schemes

Putative ligation schemes	Structures with low violations[a]	Average violations for all structures[b]	Average energies for all 20 structures[c] (kcal/mol)
6,9,28,34	9	1	518 ± 188
6,9,28,37	0	6	971 ± 388
6,9,34,37	1	4	1018 ± 336
6,28,34,37	2	5	732 ± 339
9,28,34,37	0	5	865 ± 513

Data shown from XPLOR calculations using different zinc ligation schemes as described in the text.
[a]The number of structures out of the 20 calculated for each ligation scheme which had violations of <0.5 Å and <10° is given.
[b]Violations reported are the average number of violations from all 20 structures where violations were >0.5 Å on distance restraints and >10° on angle restraints.
[c]Energies reported are the average from all 20 structures.

packing of β2 onto β1 (Figure 3C). This packing arrangement, of β1 and β2, may be important for biological function and is probably conserved in B-box domains. It is interesting to note that residues which comprise β1 (13–18) are among the most highly conserved residues within the family (Figures 1B and 3C).

A number of conserved charged residues are located on the surface of the B-box structure, in particular residue 11, which is usually a glutamic acid (Figure 1B). Also, the flexible loop region (19–25) tends to be highly charged throughout the family and also comprises two unligated Cys residues (Cys17/25). Interestingly, within this loop the B1 and B2 B-boxes have a putative metal ligand at residue 20 (Cys/His) whereas standard B-boxes have a completely conserved Asp (Figure 1B). An exception to this is ATDC B2, which resembles more a standard B-box as highlighted by the conservation of Asp at this position (Figure 1B). It is tempting to speculate that this flexible loop region forms another divalent metal-binding site which in the context of the whole XNF7 would be involved in interacting with another domain via metal ions. It is noticeable that there are some different patterns of residue conservation between the standard and the B1/B2 B-boxes which may reflect functional differences within the B-box family. However, the roles of these residues are at present unclear, and await further study.

Functional implications

Recent functional studies of XNF7 have shown that the B-box domain is probably involved in forming protein–protein interactions with components of chromosomes (Li *et al.*, 1994b) and—based on data with RFP—probably also homodimerization (T.Cao and L.Etkin, unpublished observations). Furthermore, a point mutation of His34 (HisΔAsn) in the homologous PwA33 B-box dramatically affects PwA33 association with the transcription units of lampbrush chromosomes (M.Bellini, personal communication). These data suggest that a fully folded B-box in PwA33 is important for chromosomal targeting, as His34 forms one of the core zinc ligands required for monomer B-box folding (see above).

Our modelling studies of the B-box structure and coiled-coil domains suggest that the orientation of the B-box relative to the coiled-coil is specific and probably conserved among B-box-containing proteins. For a B-box dimer model, the first β-strand β1 and parallel extended region could form a specific interaction surface on each side of the coiled-coil. Electrostatic surface potential calculations suggest that the XNF7 B-box domain is highly positive with the conformationally flexible loop negative (K.L.B.Borden, unpublished observations). It is interesting to speculate that this surface might mediate functional interactions, perhaps by extending into a β-sheet for a protein–protein interaction. In conclusion, the solution structure of the XNF7 B-box domain reported here provides a basis for in-depth site-directed mutagenesis studies to determine the molecular role of the B-box in XNF7 function and also in other B-box family members, especially the human proto-oncogene products PML, T18 and RFP.

Materials and methods

Peptide synthesis and purification

A peptide comprising residues 219–260 of XNF7 and eight subsequent mutant peptides were synthesized on a Model 431A Applied Biosystems Solid Phase Synthesizer and purified as described previously (Borden *et al.*, 1993). The yields of the 42mer peptides varied but were on average ~20%. A few micrograms of each pure peptide (peak fraction from HPLC separations) were analysed by Matrix-Assisted Laser Desorption mass spectrometry (Karas and Hillenkamp, 1988). A single species was observed with experimental molecular weights within 1 Da of the calculated masses for fully reduced forms of the peptides. The concentrations of XNF7 wild-type and mutant B-box peptides were measured optically using a calculated extinction coefficient of 0.26 for a 1 mg/ml solution measured at 280 nm. Samples were purified in the presence of ethane dithiol and in general, subsequent experiments had trace amounts of ethane dithiol present.

Optical spectroscopy and 1D ^{1}H NMR experiments

Cobalt binding studies and 1D ^{1}H NMR methods were used to determine the metal-binding properties of the XNF7 wild-type and mutant B-box peptides. The cobalt-binding reactions were monitored on an HP 8452 diode array spectrophotometer using a 1 cm path length at room temperature. The spectra were corrected by subtracting the contribution from the peptide and buffer alone. Typically, a solution (1 ml) containing 40–50 μM of XNF7 B-box peptide or mutant (10 mM Tris–HCl, pH 7.5) was titrated with a 1 mM solution of $CoCl_2$. For 1D ^{1}H NMR measurements, ~5 mg of the XNF7 B-box peptide or mutant peptide was dissolved in 10 mM sodium phosphate and 100 mM KCl at pH 7.5 (uncorrected for isotope effects) in 0.5 ml of either ^1H_2O or ^2H_2O. The proton chemical shifts were referenced to 0.1 mM internal sodium tetradimethyl-2,2,3,3-tetra-deutero-4,4-dimethyl-4-silapentanoic acid. All NMR experiments were carried out at 30°C. Argon gas was bubbled through the solution to reduce problems with precipitation. The final sample concentrations were between 1–1.5 mM. Some mutant peptides were less soluble, with final concentrations of ~500 μM. The peptides were not soluble at higher concentrations. ^1H_2O samples contained 10% ^2H_2O to provide a lock signal. When further additions of Zn^{2+} caused no further shifting of the peaks the peptide was considered to be fully saturated with metal ion.

^1H NMR experiments

Experiments were carried out at 30°C and pH ~7.5. NOESY, TOCSY, DQF-COSY and ROESY experiments were carried out in ^2H_2O and NOESY and TOCSY in ^1H_2O. 1D and 2D NMR experiments in ^2H_2O were collected at 11.7 T on a Bruker AM spectrometer or at 14.1 T on a Varian Unity spectrometer. Phase-sensitive data were collected using the method of time-proportional–phase-incrementation on the Brucker system (Marion and Wüthrich, 1983) or by the method of States *et al.* (1982) on the Varian Unity system. In all cases, sine-modulated data were recorded with proper adjustment of the receiver phase and dead time to produce a flat base-plane (Frenkiel *et al.*, 1990). A ^1H TOCSY experiment with a 50 ms mixing time using MLEV 17 (Bax and Davis, 1985) to produce the isotropic mixing was carried out. The spin-lock field strength was 7.4 kHz. The acquisition time was 170 ms in t_2 and 33.2 ms in t_1. A series of phase-sensitive NOESY experiments in ^2H_2O in which the mixing time was varied from 200–500 ms were also recorded. A total of 400 increments of 2048 points were collected with a 2.5 s recycle delay. The acquisition times were identical to those in the above TOCSY. A DQF-COSY experiment was recorded with 512 increments of 2048 points and a 2.5 s recycle delay. The acquisition time was 205 ms in t_2 and 51 ms in t_1. ROESY experiments with a 150 ms mixing time and a spin lock field strength of 3.25 kHz were also collected according to the methods of Bothner-By *et al.* (1984) and Bauer *et al.* (1990). For all of the 2D experiments carried out in ^2H_2O, the residual HDO signal was suppressed using presaturation. The free-induction decays were zero-filled to 8192 points in t_2 and 2048 points in t_1. Before Fourier transformation, experiments were apodized with a $\pi/3$ shifted sine bell squared function for Bruker data or a Gaussian function for Varian data. Both TOCSY and NOESY experiments in ^1H_2O were recorded at 14.1 T on a Varian Unity spectrometer. Both the TOCSY and NOESY experiments were collected using an echo delay to produce a flat base plane. ^1H_2O samples contained 10% ^2H_2O to provide a lock signal. The intense water signal was suppressed using presaturation. In the NOESY experiment, a 180° pulse was applied in the middle of the mixing time to further suppress the ^1H_2O signal. The NOESY experiments were collected with mixing times varied from 250–350 ms where 290 ms was determined to be the optimal mixing time. A total of 400 increments of 4096 data points were collected. The acquisition time was 294 ms in t_2 and 56 ms in t_1. A NOESY experiment (t_m = 300 ms) using a pre-TOCSY sequence in order to recover bleached α protons was also collected (Otting and Wüthrich, 1987). In this case, the experiment was collected with a pre-TOCSY mixing time of 25 ms and a spin-lock field strength of 7.7 kHz using MLEV17 to produce the isotropic mixing. The acquisition time was 256 ms in t_2 and 63 ms in t_1. The number of increments and data points are identical to those discussed above. The TOCSY echo experiment (as described above) was collected using MLEV17 to produce isotropic mixing time at a spin-lock field strength of 8.1 kHz for a mixing time of 50 ms. A total of 400 increments of 4096 data points were collected with an acquisition time of 294 ms in t_2 and 63 ms in t_1. Data were zero-filled to 8192 in t_2 and 1024 in t_1 and data processing was apodized with a Gaussian function of width 0.118 and 0.016 s shift in t_2 and 0.13 s width in t_1 and then Fourier transformed.

Generation of restraints

Standard sequential 2D ^1H NMR methods were used to assign 80% of the amide proton resonances and 91% of the side chain resonances. Pseudoatom corrections were used as described by Wüthrich (1986). A comprehensive list of NOE connectivities was compiled and divided into categories based on the intensity of the cross-peaks and then ascribed upper distance limits. Phi, Psi and Chi1 angles were also extracted from the data. Phi angles were obtained from measuring the value of the $^3J_{\alpha N}$ coupling constants in high digital resolution TOCSY experiments (Searle, 1993). Psi angles were calculated using the program PROPHET (A.N.Lane, unpublished) which uses the distances between the αH_i and NH_{i+1} and βH_1 to NH_{i+1}, and $^3J_{\alpha N}$. Chi1 values were obtained using the method of Wagner *et al.* (1987) where the distances between αH and $\beta_1 H$ and αH and $\beta_2 H$ were extracted from ROESY experiments in order to avoid the problem of spin-diffusion. Values for $J\alpha\beta_1$, $J\alpha\beta_2$, $dN\beta_1$ and $dN\beta_2$ were also used.

Structure calculations

Distance and angle constraints were used as input into XPLOR v 3.1 (Nilges *et al.*, 1988; Brünger, 1992). The distance geometry embedding procedure was carried out followed by simulated annealing, slow cooling and 1000 steps of Powell energy minimization. Unless stated otherwise, 50 structures were calculated using the XPLOR protocol with the following changes: molecules were heated to 1000 K, and were then allowed to cool slowly to 100 K in 4000 steps. From the 50 structures calculated, 39 converged into properly folded peptides while 11 formed the pseudo mirror image. The Zn^{2+} atoms were then included in the XPLOR calculations with additional constraints as previously described (Neuhaus *et al.*, 1992) to maintain the tetrahedral bonding geometry of the sites and the correct bond lengths. These structures were refined and energy minimized in XPLOR as described above and in Borden *et al.* (1995).

Acknowledgements

We thank Dr Mike Gradwell for help with XPLOR and Dr Chris Bauer for assistance with data collection. We are deeply indebted to Dr Andrew Lane for numerous discussions and encouragement throughout the project and for the program PROPHET. We also thank Drs Geoff Barton and Rob Russell for database searches. This work is supported by the Imperial Cancer Research Fund, Medical Research Council and National Science Foundation and NIH grants to L.D.E. Some peptides were synthesized by the Synthetic Antigen Laboratory at UTMDACC which is supported by NCI CA16672. Images in Figures 3 and 4 were made using the program PREPI (S.Islam and M.Sternberg, ICRF). Coordinates are being deposited in the Protein Data Bank.

References

Bauer,C.J., Frenkiel,T.A. and Lane,A.N. (1990) A comparison of the ROESY and NOESY effects for larger molecules with applications to nucleic acids. *J. Magn. Reson.*, **87**, 144–152.

Bax,A. and Davis,D.G. (1985) MLEV-17 based two dimensional homonuclear magnetisation transfer spectroscopy *J. Magn. Reson.*, **65**, 355–360.

Bellini,M., La Croix,J.-C. and Gall,J.G. (1993) A putative zinc binding protein on lampbrush chromosome loops. *EMBO J.*, **12**, 107–114.

Berg,J.M. and Merkle,D.L. (1989) On the metal ion specificity of 'zinc finger' proteins. *J. Am. Chem. Soc.*, **111**, 3759–3761.

Borden,K.L.B., Martin,S.R., O'Reilly,N., Lally,J.M., Reddy,B.A., Etkin,L.D. and Freemont,P.S. (1993) Characterisation of a novel cysteine/histidine -rich metal binding domain from *Xenopus* nuclear factor XNF7. *FEBS Lett.*, **335**, 255–260.

Borden,K.L.B., Boddy,M.N., Lally,J., O'Reilly,N.J., Martin,S., Howe,K., Solomon,E. and Freemont,P.S. (1995) The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J.*, **14**, 1532–1541.

Bothner-By,A.A., Stephens,R.L., Lee,J.-M., Warren,C.D. and Jeanloz,R.W. (1984) Structure determination of a tetrasaccharide: transient nuclear overhauser effects in the rotating frame. *J. Am. Chem. Soc.*, **106**, 811–813.

Brown,M.A. *et al.* (1994) Regulation of BRCA1. *Nature*, **372**, 733.

Brünger,A.T. (1992) *XPLOR. A System for X-Ray Crystallography and NMR.* Yale University Press, New Haven, CT.

Campbell,I.G. *et al.* (1994) A novel gene encoding a B-box protein within the BRCA1 region at 17q21.1. *Human Mol. Gen.*, **3**, 589–594.

Chan,E.K.L., Hamel,J.C., Buyon,J.P. and Tan,E.M. (1991) Molecular definition and sequence motifs of the 52 kD component of human ss-A/Ro auto-antigen. *J. Clin. Invest.*, **87**, 68–76.

de Thé,H., Lavau,C., Marchio,A., Chomienne,C., Degos,L. and Dejean,A. (1991) The PML-RARα fusion mRNA generated by the t(15;17) translocation in acute promyelocytic leukemia encodes a functionally altered RAR. *Cell*, **66**, 675–684.

Dreyer,C., Wang,Ya.H., Wedlich,D. and Hansen,P. (1983) In McClaren,A. and Wylie,C.C (eds), *Current Problems in Germ Cell Differentiation.* Cambridge University Press, Cambridge, UK, pp. 322–352.

Frankel,A.D., Bredt,D.S. and Pabo,C.O. (1988) Tat protein from human immunodeficiency virus forms a metal-linked dimer. *Science*, **240**, 70–73.

Freemont,P.S. (1993) The RING Finger: a novel protein sequence motif related to the zinc finger. *Ann. NY Acad. Sci.*, **684**, 174–192.

Freemont,P.S., Hanson,I.M. and Trowsdale,J. (1991) A novel cysteine-rich sequence motif. *Cell*, **64**, 483–484.

Frenkiel,T.A., Bauer,C.J., Carr,M.D., Birdsall,B.B. and Feeney,J. (1990) HMQC-NOESY-HMQC, a three dimensional NMR experiment which allows detection of Nuclear Overhauser effects between protons with overlapping signals. *J. Magn. Reson.*, **90**, 420–425.

Georgiadis,M.M., Komiya,H., Chakrabarti,P., Woo,D., Kornvc,J.J. and

Rees,D.C. (1992) Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii*. *Science*, **257**, 1653.

Goddard,A.D., Borrow,J., Freemont,P.S. and Solomon,E. (1991) Characterization of a novel zinc finger gene disruption by the t(15; 17) in acute promyelocytic leukemia. *Science*, **254**, 1371–1374.

Harbury,P.B., Zhang,T., Kim,P.S. and Alber,T. (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, **262**, 1401–1407.

Inoue,S. *et al.* (1993) Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proc. Natl Acad. Sci. USA*, **90**, 11117–11121.

Kakizuki,A., Miller,W.H., Umesono,K., Warrel,R.P.,Jr, Frankel,S.R., Murty,V.V.V.S., Dimetrovsky,E. and Evans,R.M. (1991) Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RARα with a novel transcription factor PML. *Cell*, **66**, 663–674.

Karas,M. and Hillenkamp,F. (1988) Laser desorption ionisation of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.*, **60**, 2299–2301.

Kastner,P., Aymee,P., Lutz,Y., Rochetty-Egly,C., Gaub,M.-P., Durand,B., Lanotte,M., Berger,R. and Chambon,P. (1992) Structure, localisation and transcriptional properties of two classes of retinoic acid receptor α fusion proteins in acute leukaemia (APL): structural similarities with a new family of oncoproteins. *EMBO J.*, **11**, 629–642.

Le Douarin,B. *et al.* (1995) The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *EMBO J.*, **14**, 2020–2033.

Leonhardt,E.A., Kapp,L.N., Young,B.R. and Murnane,J.P. (1994) Nucleotide-sequence analysis of a candidate gene for *ataxia-telangiectasia* group d(atdc). *Genomics*, **19**, 130–136.

Li,X.X., Shou,W., Kloc,M., Reddy,B.A. and Etkin,L.D. (1994a) Cytoplasmic retention of *Xenopus* nuclear factor 7 before the mid blastula transition uses a unique anchoring mechanism involving a retention domain and several phosphorylation sites. *J. Cell Biol.*, **124**, 7–17.

Li,X.X., Shou,W., Kloc,M., Reddy,B.A. and Etkin,L.D. (1994b) The association of *Xenopus* nuclear factor 7 with subcellular structures is dependent upon phosphorylation and specific domains. *Exp. Cell Res.*, **218**, 472–480.

Lovering,R. *et al.* (1993) Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc. Natl Acad. Sci. USA*, **90**, 2112–2116.

Marion,D. and Wüthrich,K. (1983) Applications of phase sensitive two dimensional correlated spectroscopy (COSY) for measurement of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, **113**, 967–974.

Michael,S.F., Kilfoil,V.J., Schmidt,M.H., Amann,B.T. and Berg,J.M. (1992) Metal binding and folding properties of a minimalist Cys_2His_2 zinc finger peptide. *Proc. Natl Acad. Sci. USA*, **89**, 4796–4800.

Miki,T., Fleming,T.P., Crescenzi,M., Molloy,C., Blam,S., Reynolds,S. and Aaronson,S. (1991) Development of a highly efficient cDNA cloning system: application to oncogene isolation. *Proc. Natl Acad. Sci. USA*, **88**, 5167–5171.

Miki,Y. *et al.* (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 66–71.

Miller,M., Kloc,M., Reddy,B.A., Eastman,E., Dreyer,C. and Etkin,L.D. (1989) x1gv7: a maternal gene product localised in nuclei of the central nervous system in *Xenopus laevis*. *Genes Dev.*, **3**, 572–583.

Miller,M., Reddy,B.A., Kloc,M., Li,X.X., Dreyer,C. and Etkin,L.D. (1991) The nuclear-cytoplasmic distribution of the *Xenopus* nuclear factor xnf-7 changes during early development coincident with its state of phosphorylation. *Development*, **113**, 569–575.

Neuhaus,D., Nakaseko,Y., Schwabe,J.W.R. and Klug,A. (1992) Solution structure of two zinc finger domains from SW15 obtained using two dimensional 1H nuclear magnetic resonance spectroscopy. *J. Mol. Biol.*, **228**, 637–651.

Nilges,M., Clore,G.M. and Gronenborn,A.M. (1988). Determination of the three-dimensional structures of proteins from inter-proton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.*, **229**, 317–324.

Otting,G. and Wüthrich,K. (1987) Pre-tocsy, a new experiment for determining 2D ^{1}H NMR of proteins in H_2O solutions. *J. Magn. Reson.*, **75** 546–549.

Patarca,R. *et al.* (1988) rpt-1, an intracellular protein from helper-inducer T cells that regulates gene expression of interleukin-2 receptor, and human immunovirus type 1 *Proc. Natl Acad. Sci. USA*, **85**, 2733–2737.

Reddy,B.A. and Etkin,L.D. (1991) A unique bipartite cysteine-histidine motif defines a sub family of potential zinc-finger proteins. *Nucleic Acids Res.*, **19**, 6330.

Reddy,B.A., Kloc,M. and Etkin,L.D. (1991) The cloning and characterization of a maternally expressed novel zinc finger nucleophosphoprotein (xnf-7). *Dev. Biol.*, **481**, 107–116.

Reddy,B.A., Etkin,L.D. and Freemont,P.S. (1992) A novel zinc finger coiled-coil domain in a family of nuclear proteins. *Trends Biochem. Sci.*, **17**, 344–345.

Russell,R.B. and Barton,G.J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins: Struct. Funct. Genet.*, **14**, 309–323.

Searle,M.S. (1993) NMR studies of drug-DNA interactions. *Prog. Nucl. Magn. Spectr.*, **25**, 403–480.

Slobbe,R.L., Pluk,W., van Wenrooij,W.J. and Pruijn,G.J.M. (1992) Ro ribonucleoprotein assembly *in vitro*. Identification of RNA–protein and protein–protein interactions. *J. Mol. Biol.*, **227**, 361–366.

States,D.J., Haberkorn,R.A. and Ruben,D.J. (1982) A two dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.*, **48**, 286–292.

Takahashi,M., Inaguma,Y., Hiai,H. and Hirose,F. (1988) Developmentally regulated expression of a human 'Finger'- containing gene encoded by the 5′ half of the *ret* transforming gene. *Mol. Cell. Biol.*, **8**, 1853–1856.

Wagner,G., Braun,W., Havel,T.F., Schauman,T., Go,N. and Wüthrich,K. (1987) Protein structure in solution by Nuclear Magnetic Resonance and Distance Geometry. The polypeptide fold of the basic pancreatic trypsin inhibitor determined using two different algorithms DISGEO and DISMAN. *J. Mol. Biol.*, **196**, 611–639.

Wishart,D.S., Sykes,B.D. and Richards,F.M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.*, **222**, 311–333.

Wüthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. John Wiley and Sons, New York.

Received on July 12, 1995; revised on September 5, 1995