

# Solution Structure of the C-terminal Domain of TFIID P44 Subunit Reveals a Novel Type of C4C4 Ring Domain Involved in Protein-Protein Interactions\*

Received for publication, November 17, 2004, and in revised form, March 15, 2005  
Published, JBC Papers in Press, March 24, 2005, DOI 10.1074/jbc.M412999200

Esther Kellenberger<sup>‡§</sup>, Cyril Dominguez<sup>¶</sup>, Sébastien Fribourg<sup>‡</sup>, Emeric Wasielewski<sup>‡</sup>,  
Dino Moras<sup>‡</sup>, Arnaud Poterszman<sup>‡</sup>, Rolf Boelens<sup>||</sup>, and Bruno Kieffer<sup>‡\*\*</sup>

From the <sup>‡</sup>Département de Biologie et de Génomique Structurales, Institut de Génétique et de Biologie Moléculaire et Cellulaire UMR 7104, 1, rue Laurent Fries, BP 10142, 67404 Illkirch cedex, France and <sup>||</sup>Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

**The human general transcription factor TFIID is involved in both transcription and DNA nucleotide excision repair. Among the 10 subunits of the complex, p44 subunit plays a crucial role in both mechanisms. Its N-terminal domain interacts with the XPD helicase, whereas its C-terminal domain is involved specifically in the promoter escape activity. By mutating an exposed and non-conserved cysteine residue into a serine, we produced a soluble mutant of p44-(321–395) suitable for solution structure determination. The domain adopts a C4C4 RING domain structure with sequential organization of  $\beta$ -strands that is related to canonical RING domains by a circular permutation of the  $\beta$ -sheet elements. Analysis of the molecular surface and mutagenesis experiments suggests that the binding of p44-(321–395) to TFIID p34 subunit is not mediated by electrostatic interactions and, thus, differs from previously reported interaction mechanisms involving RING domains.**

TFIID is a multiprotein complex that is required both in DNA repair through the nucleotide excision repair pathway and transcription (1, 2). Mutations in TFIID subunits lead to severe genetic diseases such as xeroderma pigmentosum, Cockayne's syndrome, and trichothiodystrophy (3, 4). TFIID contains nine large subunits that form two structural and functional complexes; "core" TFIID, comprising subunits p52, p34, p62, p44, and the XPB helicase, and the Cdk-activating kinase complex, comprising cdk7 kinase, MAT1, and cyclin H (Fig. 1A). The ninth subunit, XPD helicase, is associated with either the core or the Cdk-activating kinase complex. Recently, an additional low molecular mass (8 kDa) subunit of TFIID was identified, and its role in the DNA repair syndrome trichothiodystrophy group A was described (5). Several studies aiming at deciphering structure-activity relationships for components of

the TFIID complex have recently been reported. The structures of cyclin H (8) and domains of MAT1 (9) and p62 (10) have been solved at atomic resolution, and lower resolution electronic microscopy studies of the entire complex revealed a ring-like structure (6, 7).

Among the core TFIID subunits, p44 plays a central role in the transcription/repair activities of TFIID with distinct functions associated with each of its three domains (Fig. 1B); the N-terminal domain specifically regulates the DNA repair activity of the XPD helicase subunit (11), the central domain is essential for p62 subunit incorporation within TFIID, and the C-terminal domain is mainly involved in promoter escape activity after the synthesis of the first phosphodiester bond by RNA polymerase II (12) but also interacts strongly with another core subunit, p34 (13). Mass spectrometry analysis showed that p44 binds three zinc ions, one by a C4 zinc finger motif in the central domain (residues 252–320) and two by the C-terminal domain (14). The sequence of the C-terminal domain of p44 reveals a high content of cysteine and histidine residues. The preliminary structural characterization of the C-terminal domain showed that the two zinc ions are bound using a cross-braced arrangement of coordinating residues in the sequence similar to RING domains. Identification of the eight coordinating residues by NMR was severely hindered by line broadening due to oxidative processes and intermediate time-scale dynamics as well as the high number of conserved putative zinc-coordinating residues. More especially, NMR data recorded on the wild-type p44-(321–395) domain were compatible with both a C6H2 and a C4C4 pattern of zinc coordinating residues, preventing a complete description of the domain structure (Fig. 1C). To determine unambiguously the zinc binding mode of the C-terminal domain of p44 and gain insight into its role within the TFIID core particle, we expressed and purified the C381S mutant of this domain. The mutation of this non-conserved, solvent-exposed cysteine residue led to a dramatic improvement in the behavior of the domain, allowing us to define the zinc binding motif and to report a refined solution structure for this domain. Mutations of conserved hydrophobic residues on the surface of the domain provided further structural insights into the interaction with p34, unraveling a specific functional aspect of this new variant of RING domains.

## MATERIALS AND METHODS

**Sample Preparation**—The C381S mutant of p44-(321–395) was cloned into a modified version of pGEX-4T2 (Amersham Biosciences) as previously described for other p44 mutants (14). Unlabeled and <sup>15</sup>N-labeled recombinant proteins were expressed in *Escherichia coli* BL21(DE3) as glutathione *S*-transferase fusion protein using LB me-

\* This work was supported by funds from European Community program "Access to large scale facility," Université Louis Pasteur de Strasbourg, the CNRS, the INSERM, the Netherlands Foundation for Chemical Research (NWO-CW), the Center for Biomedical Genetics, and SPINE Contract QL2-CT-2002-00988. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: UMR 7081, Bioinformatics group, Faculté de Pharmacie, 74 route du Rhin BP24, 67401 Illkirch Cedex, France.

¶ Present address: Institute for Molecular Biology and Biophysics, Swiss Federal Institute of Technology Zurich, ETH-Honggerberg, CH-8093 Zurich, Switzerland.

\*\* To whom correspondence should be addressed. kieffer@esbs.u-strasbg.fr.

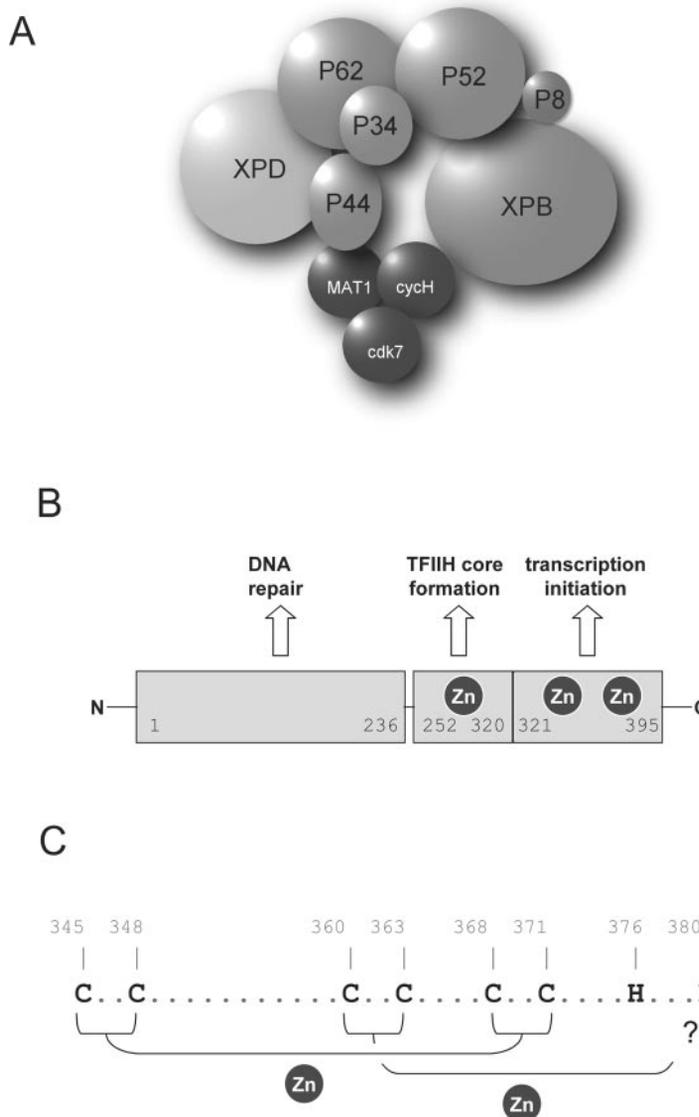


FIG. 1. *A*, molecular composition of TFIIH. Subunits in *dark gray* belong to the Cdk-activating kinase complex, and those in *light gray* belong to the TFIIH core complex. The XPD helicase is found associated with either the Cdk-activating kinase complex or core complex. *B*, domain organization of the p44 subunit with associated zinc content and biological functions. The *numbering* is based on the human protein sequence. *C*, pattern of conserved putative zinc-coordinating residues in p44-(321–395).

dium and  $^{15}\text{N}$ -labeled rich medium (Silantes GmbH, Germany), respectively. Both media were supplemented with  $6\ \mu\text{M}\ \text{ZnCl}_2$ . Transformed bacteria were grown at  $37^\circ\text{C}$ . Protein expression was induced at  $A_{600} = 0.65$  with  $1\ \text{mM}$  isopropyl- $\beta$ -D-thiogalactopyranoside and harvested by centrifugation after 4 h at  $30^\circ\text{C}$ . Cells were suspended in lysis buffer ( $50\ \text{mM}$  Tris-HCl (pH 8),  $200\ \text{mM}$  NaCl,  $10\%$  glycerol,  $0.1\%$  Triton X-100,  $5\ \text{mM}$   $\beta$ -mercaptoethanol,  $0.1\ \text{mg/ml}$  lysozyme, EDTA-free protease inhibitor mix), disrupted by sonication, and centrifuged at  $100,000 \times g$  for 1 h at  $4^\circ\text{C}$ . The protein was purified from the cleared lysate using a GSH-Sepharose column (Amersham Biosciences), and the glutathione *S*-transferase was removed using bovine thrombin, leaving additional residues (glycine-serine-histidine) at the N terminus of the desired protein before the methionine. p44-(321–395) C381S mutant protein was finally purified using a Superdex 75 gel filtration column.

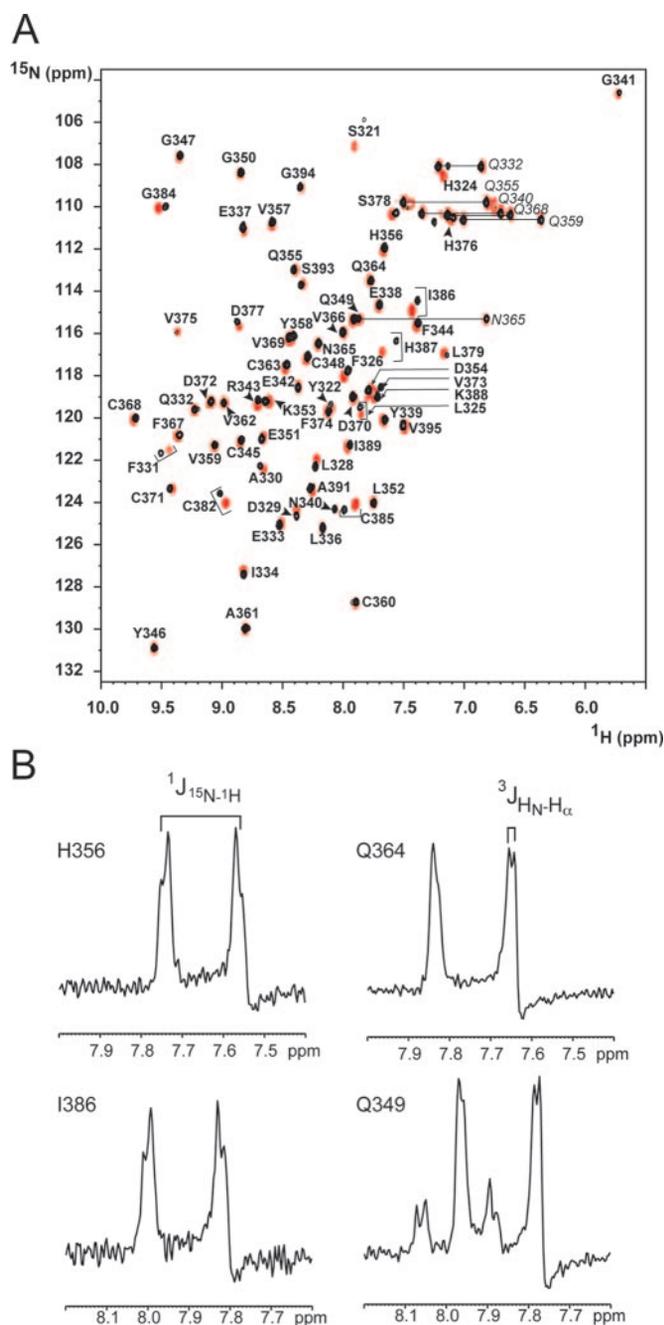
$0.5\ \text{mM}$  unlabeled and  $1\ \text{mM}$   $^{15}\text{N}$ -labeled proteins were dissolved in  $20\ \text{mM}$  deuterated Tris-HCl buffer,  $20\ \text{mM}$  NaCl, and  $0.5\ \text{mM}$  dithiothreitol at pH 7 (at  $4^\circ\text{C}$ ).  $^{113}\text{Cd}$ -loaded protein was obtained by adding  $^{113}\text{Cd}$ -EDTA at a final concentration of  $4\ \text{mM}$  to the  $^{15}\text{N}$ -labeled zinc-loaded protein, and the kinetics of the cadmium/zinc substitution process were followed by recording a series  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC<sup>1</sup> spectra. All samples were kept at  $4^\circ\text{C}$  under argon. For periods exceeding 1 month, proteins were stored at  $-80^\circ\text{C}$ .

**p44-p34 Interaction Experiments**—The cDNA encoding regions of human p34-(1–242) and p44-(321–395) were cloned into the expression vectors pET-15b (N-terminal His tag, ampicillin resistance, ColE1 ori-

gin) and pACYC184–11b (no tag, chloramphenicol resistance, and p15A origin), respectively. Both vectors harbor compatible replication origins suitable for co-expression of the two subunits as described in Fribourg *et al.* (13). The mutations C345A, F344S, F344S/Y346S, V366S, D370R/D372R, D370R, and D372R in p44-(321–395) were obtained by PCR-assisted mutagenesis using Deep Vent DNA polymerase (Biolabs) and the appropriate oligonucleotides. For co-expression experiments, BL21  $\lambda$ DE (3) bacteria were electroporated with a mix of  $50\ \text{ng}$  of pET-15b and pACYC184–11b constructs and grown on Petri plates containing  $200\ \mu\text{g/ml}$  ampicillin and  $35\ \mu\text{g/ml}$  chloramphenicol. Cultures were grown from single colonies in  $10\ \text{ml}$  of LB medium at  $37^\circ\text{C}$  and induced at  $A_{600} = 0.5$  by the addition of isopropyl- $\beta$ -D-thiogalactopyranoside to a final concentration of  $0.6\ \text{mM}$ . After 3 h at  $25^\circ\text{C}$  the bacteria were collected by centrifugation. Pellets were resuspended in  $1\ \text{ml}$  of  $50\ \text{mM}$  Tris-HCl, pH 8,  $400\ \text{mM}$  KCl (buffer A) with  $5\ \text{mM}$   $\beta$ -mercaptoethanol, sonicated, and centrifuged ( $13,000\ \text{rpm}$ ,  $4^\circ\text{C}$ ) for 30 min in Eppendorf tubes. Clear lysates were incubated with  $50\ \mu\text{l}$  of cobalt affinity resin (Clontech<sup>TM</sup>) and extensively washed with buffer A. Bound proteins were analyzed by SDS-PAGE using Coomassie Blue staining.

**NMR Experiments**—NMR measurements were carried out at 293 and 303 K on Bruker DRX-600 spectrometers. The assignment of the C381S mutant of p44-(321–395) was based on the assignment of the wild-type domain and  $^{15}\text{N}$ -edited two-dimensional total correlation and two-dimensional NOE spectroscopy spectra for the regions affected by the mutation. A  $^1\text{H}$ ,  $^{113}\text{Cd}$  HSQC spectrum was recorded using a broadband *z*-gradient probe at 293 K. The delay for magnetization transfer ( $1/2J_{\text{H-}^{113}\text{Cd}}$ ) was set to 12 ms.  $256\ t_1$  increments of 128 scans were recorded giving a total length for the experiment of 11 h. The  $^{113}\text{Cd}$  chemical shifts are reported relative to  $1\ \text{M}\ \text{CdSO}_4$ . The protonation

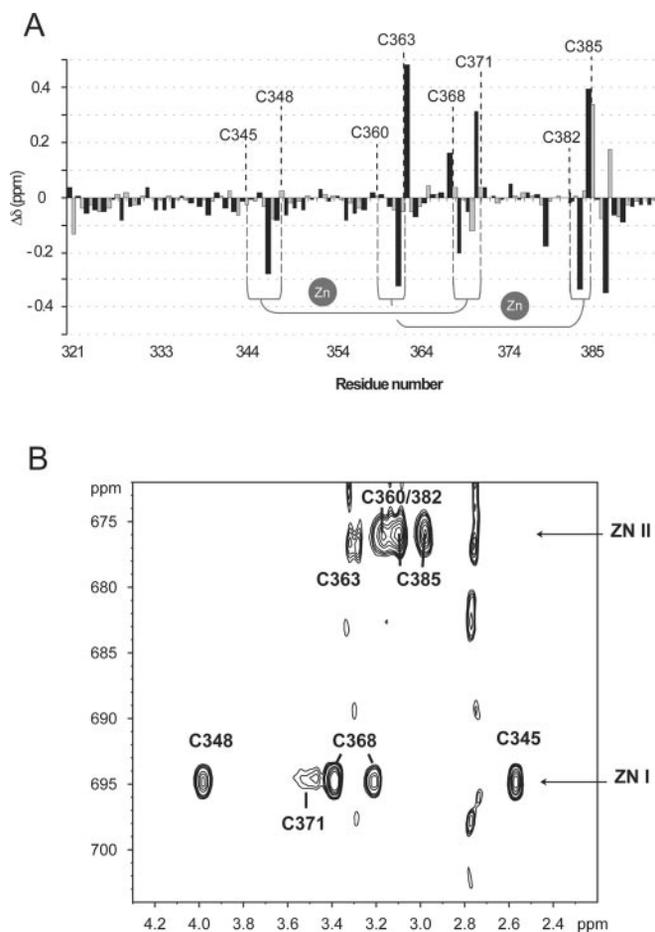
<sup>1</sup> The abbreviations used are: HSQC, heteronuclear single quantum correlation; NOE, nuclear Overhauser effect; r.m.s.d., root mean square deviation; E2, ubiquitin-conjugating enzyme.



**FIG. 2.  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra of p44-(321–395) domain at 303 K.** A, superposition of spectra of  $^{15}\text{N}$ -labeled wild-type (*black*) and mutant (*red*) proteins. Side-chain amide resonances are identified by labels in *italics*. B, cross-section along the F2 dimension through selected amide proton resonances in the  $^{15}\text{N}$ -coupled  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum.

state of the imidazole rings of histidine residues was investigated using long range  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectra recorded on the zinc- and cadmium-loaded proteins with a magnetization transfer delay ( $1/2J_{\text{H-N}}^{15\text{N}}$ ) of 23 ms. The protonation state was deduced from the pattern of observed correlations as described (15). Vicinal scalar  $^3J_{\text{HN-H}\alpha}$  coupling constants were measured from the ratio of diagonal and cross-peak intensities in a three-dimensional HNHA experiment and converted into  $\varphi$ -dihedral angle restraints (16).

**Structure Calculation**—The assignments of  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts and measurements of cross-peak intensities were achieved using XEASY (17). The NMR structure calculations were performed in CNS (18) using ARIA1.2 using default values for lower and upper distance restraints and for acceptance threshold (19). The automated assignment of the NOE cross-peaks was based on chemical shifts of  $^1\text{H}$  nuclei, allowing a frequency deviation of  $\pm 0.015$  ppm in the acquisition dimension and  $\pm 0.025$  ppm in the indirect dimension. Structure calculations were started from a random initial structure that included the two zinc



**FIG. 3. Zinc-cadmium replacement experiment.** A, chemical shift differences of the amide  $^1\text{H}$  (*black bars*) and  $^{15}\text{N}$  (*gray bars*) nuclei between the zinc- and cadmium-loaded proteins. The values of  $\Delta\delta$  were calculated as follows:  $\Delta\delta = \delta(\text{Zn}^{2+} \text{ form}) - \delta(\text{Cd}^{2+} \text{ form})$ . B,  $^1\text{H}$ ,  $^{13}\text{C}$  HSQC spectrum of the cadmium-loaded sample.

atoms assuming a tetrahedral geometry of the zinc coordination (Zn-S $\gamma$  bond lengths were set to 2.3 Å, and C $\beta$ -S $\gamma$ -Zn and S $\gamma$ -Zn-S $\gamma$  angles were set to 109.5°). Of the 50 structures computed in the final iteration, the 10 lowest energy structures were refined in a shell of water molecules using a torsion angle data base potential (20) using XPLOR-NIH (21). The final ensemble of conformations was analyzed by PROCHECK (22) and visualized with MOLMOL (23). Structural alignments were obtained using WHATIF (24).

## RESULTS AND DISCUSSION

**The Fold of the C381S Mutant of p44-(321–395) Is Identical to That of the Wild-type Domain**—The ability of the 75 C-terminal residues of the human TFIIF p44 subunit to bind two zinc atoms and fold autonomously into a RING domain-like structure was reported previously (14). However, extensive line broadening observed for several residues prohibited the precise definition of the second zinc binding motif despite extensive mutagenesis of putative zinc coordinating residues. Assuming this behavior of the p44-(321–395) domain to result from possible oxidative processes involving exposed cysteines, we mutated cysteine 381, which is not conserved among species, into a serine residue. The mutation led to a striking improvement in behavior of the recombinant protein. The sample lifetime was dramatically lengthened as shown by the reproducibility of NMR experiments over time. Moreover, values of translational diffusion coefficients measured using pulsed field gradient experiments were compatible with diffusion as a monomer (data not shown), suggesting that, in contrast to the wild-type protein, the C381S mutant of p44-(321–395) is not prone to oxidation.

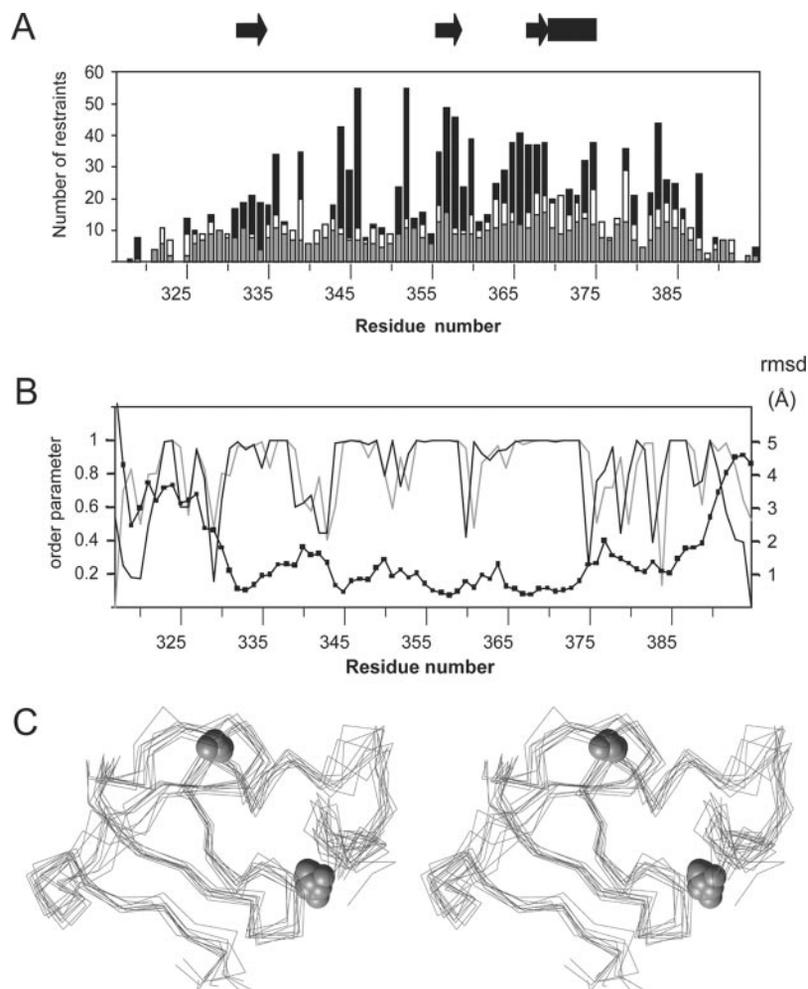


FIG. 4. **Solution structure of p44-(321–395).** *A*, number of NOE-derived restraints. Sequential, medium range, and long range NOE are indicated by *gray*, *white*, and *black bars*, respectively. Secondary structure elements are depicted on the *top*. *B*, angular order parameters and r.m.s.d. values for backbone heavy atoms as a function of sequence. Order parameters calculated for  $\phi$  and  $\psi$  dihedral angles are indicated by a *gray* and a *black solid line*, respectively. *C*, superposition of the 10 best structures obtained after the final ARIA iteration. The superposition was performed using the  $C\alpha$  atoms of residues 328–386. The positions of the zinc ions are shown by *light gray spheres*.

Chemical modification of wild-type p44-(321–395) showed that Cys-381 is the only solvent-accessible cysteine (data not shown). The three-dimensional structure of the C381S mutant of p44-(321–395) was, therefore, expected to be very similar to that of the wild-type protein. Indeed, the pattern of cross-peaks in a  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum of the mutant closely matches that of the wild-type protein, indicating that the mutation does not alter the fold of the domain (Fig. 2A). Sequence-specific assignments of the C381S mutant of p44-(321–395) were obtained from both homonuclear and  $^{15}\text{N}$ -edited spectra.  $^1\text{H}$ ,  $^{15}\text{N}$  correlations for 8 residues in the N-terminal region (including the four additional residues, GHSM, resulting from cleavage of the glutathione *S*-transferase fusion protein) were not observed, and the resonances for those residues were, therefore, left unassigned. An analysis of chemical shift differences between wild-type and mutant p44-(321–395) along the protein backbone confirms that the mutation does not lead to any major modification of the protein structure; chemical shift differences for non-exchangeable protons never exceeds 0.1 ppm, even for residue 381. The mutation mainly affects the chemical shifts of a small continuous stretch of residues C-terminal to the mutation site (Cys-382 to His-387). A few residues in the N-terminal part of the domain (Ser-321 to Phe-331) also undergo chemical shift modification upon mutation despite their remote location in the sequence. These changes may indicate that the N-terminal region, whose conformation could not be determined for both the wild-type and mutant proteins, interacts with residues in the vicinity of the second zinc-binding site.

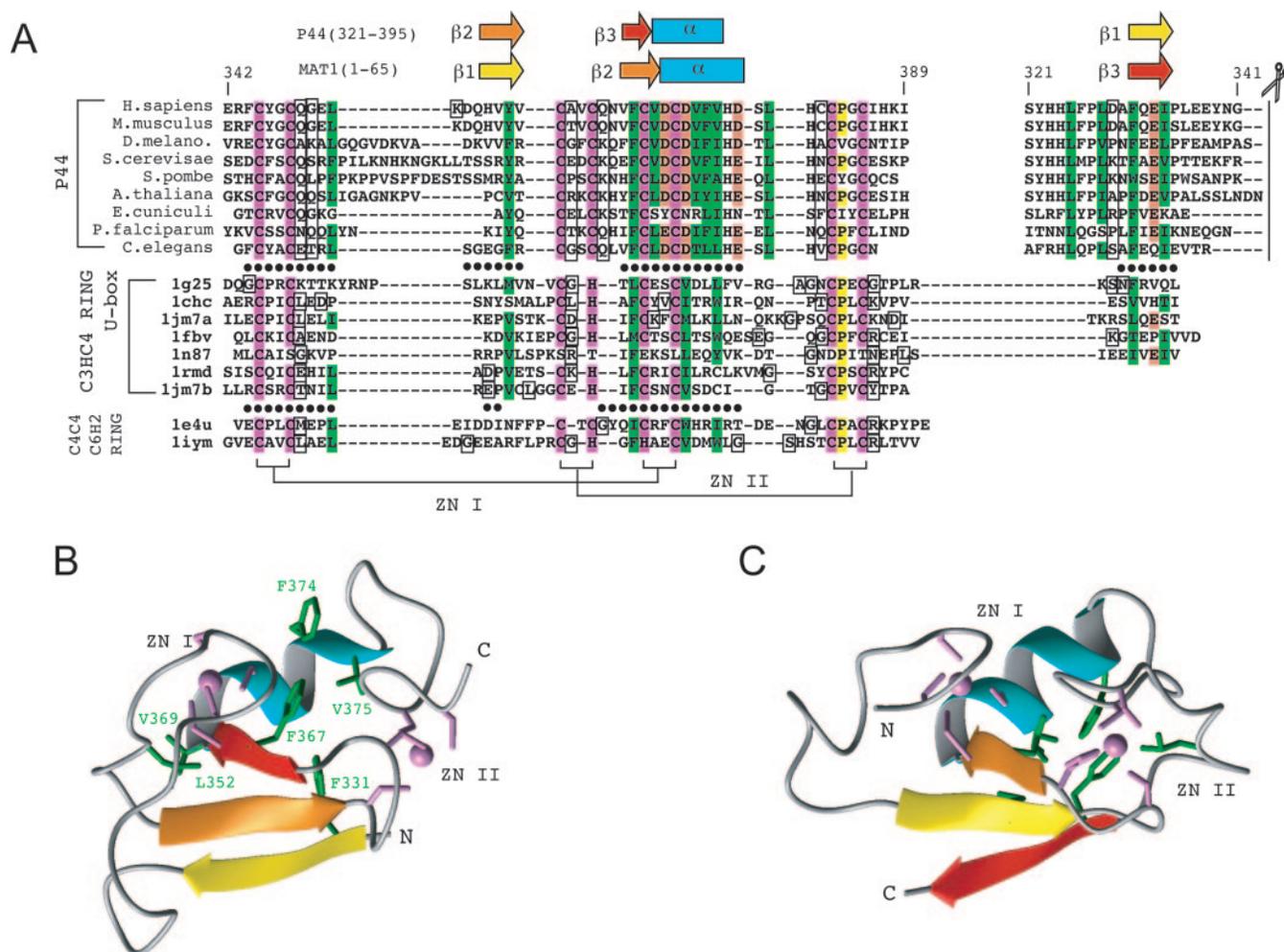
The C381S mutation results in a global and a site-specific reduction in line width for several resonances in the  $^1\text{H}$ ,  $^{15}\text{N}$

TABLE I  
Structural statistics for p44-(321–395)

	C381S mutant	Wild type
Experimental input		
Total unambiguous NOE	1294	1082
Intraresidue	525	451
Sequential	307	271
Medium range	135	111
Long range	327	249
Total hydrogen bonds restraints	17	12
Torsion angle constraints	32	25
Dihedral $\phi$	16	12
Dihedral $\psi$	16	13
Procheck statistics on 328–386 fragments <sup>a</sup>		
Residues in most favored regions	81.0 (46.0)	33.3
Residues in allowed regions	11.2 (39.4)	31.4
Residues in generously allowed regions	6.9 (10.0)	31.4
Residues in disallowed regions	1.0 (4.6)	3.9
r.m.s.d. of $C\alpha$ atoms (residues 328–386) <sup>b</sup>	0.98	1.79
r.m.s.d. of all carbon atoms (residues 328–386)	1.01	1.97

<sup>a</sup> Numbers in parentheses indicate the percentages obtained from a set of structures calculated without using conformational database to allow the comparison with the set previously obtained for the wild-type protein.

<sup>b</sup> r.m.s.d., deviation between the ensemble of structures and the mean structure in Å.

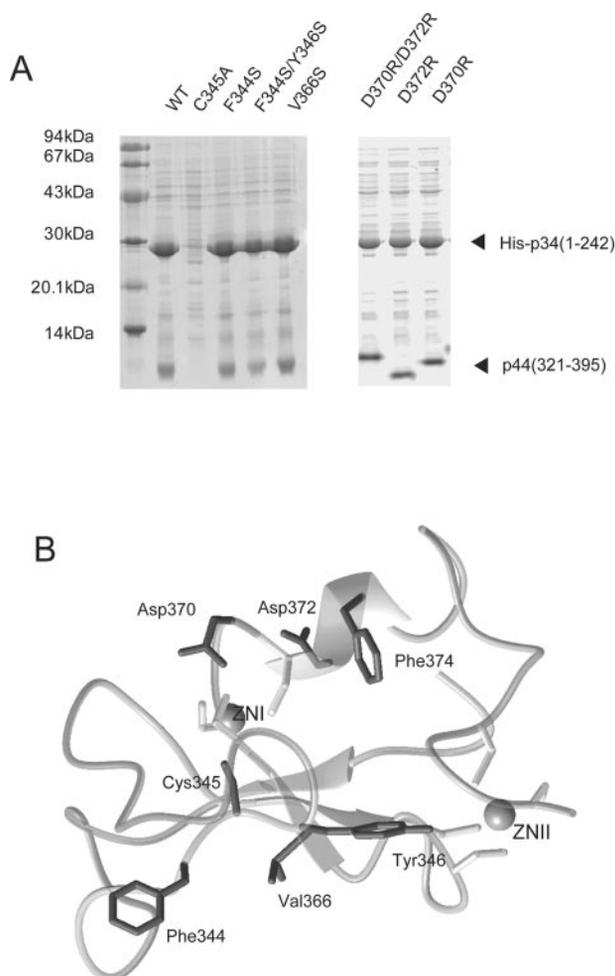


**FIG. 5. Comparison of p44-(321–395) with RING domains.** **A**, alignment of p44-(321–395) sequences and structural alignment of RING domains from the FSSP data base. *Filled circles* indicate structurally equivalent positions (r.m.s.d. less than 2.5 Å) between the RING domains of p44-(321–395) and RAG1 (PDB code 1g25) and between p44-(321–395) and the C4C4 RING domain of CNOT4 (PDB code 1e4u). The metal-coordinating residues are shown in *magenta*. Conserved hydrophobic residues are shown in *green*, and conserved positive residues are shown in *orange*. *Boxed positions* indicate residues that have positive  $\phi$  angles in the C381S mutant of the human p44-(321–395) sequence. The structural alignment of p44-(321–395) sequences with those of other RING domains requires the N-terminal of p44-(321–395) to be cut and pasted at the end of the p44 RING domain at a position indicated by a *pair of scissors*. The secondary structure of p44 and MAT1 RING domains is indicated above the alignment using a color scheme that follows the order in the primary sequence (from *yellow* for  $\beta$ -1 to *red* for  $\beta$ -3). Shown are ribbon diagrams of the C4C4 RING domain, p44-(321–395) (**B**) and of the C3HC4 RING domain, MAT1(1–65) (**C**). For the sake of clarity, only residues 328–387 have been represented for p44. Helices are shown in *blue*, and the  $\beta$ -sheets are colored as described above. The side chains of conserved hydrophobic residues are shown in *green*.

HSQC spectrum, indicating a decrease of both the overall correlation time and the exchange contributions of  $\mu$ s-ms time scale motions. This effect is particularly striking for 2 residues preceding the mutation (Val-375 and Leu-379) (Fig. 2A). The overall reduction in line widths allowed the observation of the homonuclear proton coupling multiplet patterns for most of the peaks in a  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum recorded without  $^{15}\text{N}$  decoupling. The relative intensities of inner and outer components of the multiplet are modulated by cross-correlated relaxation of HN-N and HN-H $\alpha$  dipolar interactions and, therefore, provide an efficient way to determine the sign of  $\phi$  angles in medium-sized proteins (25). The pattern observed for both Gln-364 and Gln-349 indicates positive values of  $\phi$  angle for these residues (Fig. 2B).

**The Two Zinc Ions of p44-(321–395) Are Coordinated by Eight Cysteine Residues**—Unambiguous assignment of the zinc binding residues of the C381S mutant of p44-(321–395) was achieved by slow replacement of zinc ions by cadmium ions (26). Shortly after the addition of a 4-fold excess of  $^{113}\text{Cd}^{2+}$ -EDTA to the protein sample, several cross-peaks in the  $^1\text{H}$ ,  $^{15}\text{N}$  HSQC spectrum decreased in intensity, whereas a number of

new peaks appear in the spectra, indicative of slow exchange between zinc- and cadmium-loaded proteins. Complete replacement of zinc ions by cadmium ions was achieved after 6 h, after which a single set of resonances was observed. Only a small number of resonances were shifted upon ion exchange, indicating that the exchange did not alter the structure of the domain (Fig. 3A). Chemical shift differences are clustered in four distinct segments of the protein sequence. The first three segments encompass the first three pairs of cysteine residues (Cys-345—Cys-348, Cys-360—Cys-363, Cys-368—Cys-371) that were previously identified as involved in zinc coordination (14). The fourth segment includes Cys-382 and Cys-385, the last pair of putative zinc coordinating residues providing direct evidence for the involvement of these residues in the second zinc-binding site. The largest chemical shift changes are observed for Gly-347, Val-362, Cys-363, Cys-371, Gly-384, Cys-385, and His-387, which are located at position  $i+2$  or  $i+3$  relative to a zinc-coordinating residue. These large shifts may be due to the presence of hydrogen bonds between the amide proton of these residues and the sulfur atom of the preceding cysteine residue (27). A  $^1\text{H}$ ,  $^{113}\text{Cd}$  HSQC spectrum of  $^{113}\text{Cd}$ -



**FIG. 6. Interaction of p44-(321-395) with p34.** *A*, profile of p34/p44 co-expression. The binding of mutants of p44-(321-395) to p34(1-242) is assessed by obtaining soluble His-tagged p34 protein when the two proteins are co-expressed in *E. coli*. Affinity-purified complexes were separated on a denaturing SDS-PAGE gels. *B*, ribbon diagram of p44-(321-395) highlighting mutation sites. Side chains of mutated residues are shown in *black*.

loaded C381S mutant of p44-(321-395) was recorded to unambiguously establish the zinc coordinating pattern of p44-(321-395) (Fig. 3*B*). The chemical shift values of cadmium resonances, 676.5 and 695 ppm, are consistent with those of a  $\text{Cd}^{2+}$  ion coordinated by four sulfur ligands (28). The correlations between the two cadmium resonances and the  $^1\text{H}\beta$  resonances of the coordinating residues show that Cys-345, Cys-348, Cys-368, and Cys-371 (site 1) bind one metal ion resonating at 695 ppm and that Cys-360, Cys-363, Cys-382, and Cys-385 (site 2) bind the other metal ion at 676.5 ppm. It was suggested previously that His-376 and His-380, residues conserved among p44 sequences, might be involved in zinc coordination. To probe their possible role in metal binding, long range  $^1\text{H},^{15}\text{N}$  HSQC were recorded on the zinc- and cadmium-loaded C381S mutant of p44-(321-395) (data not shown). No significant shifts for resonances of the proton and nitrogen nuclei of the imidazole ring were observed upon ion exchange, suggesting that histidine residues are distant from the metal-binding sites. In addition, no  $^{113}\text{Cd},^{15}\text{N}$  couplings were observed in long range  $^1\text{H},^{15}\text{N}$  HSQC spectrum recorded on the cadmium-loaded sample. These results provide clear evidence that p44-(321-395) binds two zinc atoms that are coordinated by 8 cysteine residues in a cross-braced manner, similar to that observed in canonical RING fingers.

**Solution Structure of the C381S Mutant of p44-(321-395)**—The sharp reduction in line broadening due to exchange contributions allowed the solution structure of p44-(321-395) to be refined. The dataset for structure calculations consisted of 1294 unambiguous intra- and interresidue distance restraints derived from two-dimensional NOE spectroscopy spectra, 17 hydrogen bond distance restraints, and 32 backbone dihedral restraints including positive  $\phi$  angles determined from  $^1\text{H}$ -coupled  $^1\text{H},^{15}\text{N}$  HSQC spectra (Fig. 4*A* and Table I). For the final ensemble of the 10 lowest energy structures resulting from ARIA calculations, the root mean square deviation for the  $\text{C}\alpha$  positions of residues 328–386 is 0.98 Å, indicating a significant improvement in the quality of the structures (Fig. 4*B*, Table I). A significant increase in the number of long range NOEs in the C-terminal part of the domain (residues 380–387) (Fig. 4*A*) allowed the definition of a region whose conformation could not previously be determined due to extensive line broadening in the wild-type protein. This is reflected in lower r.m.s.d. values for the C-terminal part of the domain, encompassing the last pair of cysteine residues involved in the second zinc-binding site (Cys-382, Cys-385) (Fig. 4*C*). The central region of p44-(321-395) spanning residues 328 and 388 is well structured, whereas the structure of the N- and C-terminal regions is less precisely defined due to a lack of experimental restraints.  $^{15}\text{N}$  heteronuclear relaxation measurements show that the corresponding amide groups are affected by motions on multiple time scales, preventing the definition of a single conformation for these regions (data not shown). The improvement in the quality of the structure is readily apparent from the Ramachandran plot statistics. Positive  $\phi$  angles were experimentally defined from cross-correlated relaxation rates for Gln-349 and Gln-364, accounting for more than half of the residues in the disallowed regions. Other residues whose dihedral angles lie in less favored regions are located mostly in poorly defined regions of the structure and display high values of the  $\phi$  angle order parameter (Fig. 4*B*).

The solution structure of the C381S mutant of p44-(321-395) is similar to that of the wild-type (Fig. 4*C*), with an r.m.s.d. of 1.9 Å resulting from the best-fit superimposition of 33  $\text{C}\alpha$  atoms in the well defined parts of the average structures. The structure consists of a triple-stranded anti-parallel  $\beta$ -sheet comprising residues Phe-331—Ile-334 ( $\beta$ 1), Gln-355—Val-359 ( $\beta$ 2), and Val-366—Cys-368 ( $\beta$ 3) followed by a short C-terminal  $\alpha$ -helix that is tightly packed onto the  $\beta$ -sheet. Both zinc ions are involved in stabilizing loops, one connecting the first and the second strand of the  $\beta$ -sheet (loop1, Pro-335 to Asp-354) and the other running from the end of the  $\alpha$ -helix to the C terminus (loop2, His-376 to Lys-388). The distance between the two zinc atoms is  $14.8 \pm 0.5$  Å. The conformation of the N-terminal part of loop1 between Pro-335 and the first zinc binding residue Cys-345 is poorly defined due to high flexibility.

**Structural Relationship with Other RING/U-box Domains**—The fold of p44-(321-395) is stabilized by a cluster of conserved hydrophobic residues including Phe-331, Leu-352, Tyr-358, Phe-367, and Val-375, and a similarity with the C3HC4 RING domains was noted previously (14). The structural alignment of p44-(321-395) sequences with a similar analysis of C3HC4 RING and U-box domains from the FSSP data base (29) shows that these hydrophobic residues are conserved across the whole family of RING folds (Fig. 5*A*). Interestingly, four of these residues are also conserved in the U-box domains. U-box domains have a similar structure to that of RING domains, but they lack the two zinc ions whose contribution to the stability of the fold is replaced by an extensive set of hydrogen bonds (30, 31). It is worth noting that the structural superposition of p44-(321-395) on a triple-stranded C3HC4 RING domain such

as MAT1(1–65) requires a circular permutation of  $\beta$  strands (Fig. 5). Strands  $\beta 1$ ,  $\beta 2$ , and  $\beta 3$  of p44-(321–395) correspond to strands  $\beta 2$ ,  $\beta 3$ , and  $\beta 1$  of the MAT1 RING domain, respectively. Structural alignment shows that the structural similarity is mostly localized in the  $\beta 2$ - $\beta 3$  hairpin region, the  $\alpha$ -helix and the first zinc-binding site, whereas a greater divergence is observed for the second site. An interesting structural feature that emerges from the comparison of available RING domain structures is the systematic presence of residues with positive  $\phi$  angles after cysteine residues (Fig. 5A). This feature was observed previously for zinc fingers (32) and is due to a particular conformation of the  $\beta$  hairpin loop that hosts a pair of zinc-coordinating residues. It appears to be a characteristic of RING domains (33).

The pattern of zinc-coordinating residues in p44-(321–395) is reminiscent of the C4C4 RING domain of CNOT4, a component of the CCR4-NOT complex that is also involved in the transcription process (26). The two domains share a characteristic spacing of four residues between the second and third pairs of zinc-coordinating residues. Although the CNOT4 RING domain lacks standard  $\beta$  strands of p44-(321–395), a reasonable superposition is possible for residues in the vicinity of the first zinc-binding site and in the  $\beta 2$  and  $\beta 3$  strands of p44-(321–395).

**Biological Relevance of the p44-(321–395) RING Structure**—It has been suggested that many RING domains are involved in regulating protein levels by mediating ubiquitin-conjugating enzyme (E2)-dependent ubiquitination (34). This function has been experimentally demonstrated for a number of RING domains and also for the U-box domain (35, 36). For CNOT4, the specific involvement of the C4C4 RING domain as ubiquitin-protein ligase has been demonstrated biochemically, and its interaction with the ubiquitin-conjugating enzyme (E2) UbcH5B has been described in detail (37, 38). Site-directed mutagenesis established the importance of negatively charged residues for the interaction (39). A comparison of p44 sequences with those of other RING domains reveals specific features of the p44 C4C4 RING (Fig. 5A). Its  $\alpha$ -helix contains two conserved aspartic acid residues and higher number of hydrophobic residues than other RING domains. It should be noted that the position of Asp-372, which follows the third pair of zinc binding residues, is occupied by a hydrophobic residue in all other RING structures. These sequence features may reflect a distinct biological function of the C4C4 RING domain of p44. No interaction with enzymes from the ubiquitination pathway has yet been reported so far for TFIIF subunits including p44. However, p44 interacts strongly with several other subunits of the TFIIF complex, including p62 and the two helicases XPB and XPD as well as p34 (40, 41). For this latter interaction, it was further shown that co-expression of p44-(321–395) with the N-terminal 242 residues of p34 produces soluble complexes, emphasizing the strength of this interaction (13). To probe the possible role of certain residues in the interaction with p34, point mutations were introduced into p44-(321–395), mapping the first zinc-binding site and the  $\alpha$ -helix. The mutation of hydrophobic residues in the vicinity of the first zinc-binding site does not affect binding to p34, but the mutation of a single cysteine residue forming this site leads to the loss of a soluble complex, showing that the conservation of the fold is required for binding. In the  $\alpha$ -helix, the mutation of the conserved acidic residues Asp-370 and Asp-372 into arginine does not prevent binding to co-expressed p34-(1–242), suggesting that the formation of the complex is not governed by electrostatic interactions involving these residues (Fig. 6). A striking feature of the p44-(321–395)  $\alpha$ -helix is the presence of Phe-374, which is exposed to the solvent. In RING domains, which are involved in the ubiquitination pathway, this residue is replaced by a polar

residue, which is important for binding to the E2-conjugating enzyme (37, 42). Mutation of Phe-374 into serine results in no stable complex being formed. Although the lower expression rate of soluble p44-(321–395) for this mutant does not allow unambiguous conclusions to be drawn, it appears that binding to p34 is not mediated by electrostatic interactions but more likely by hydrophobic contacts, thus differing from the interaction mechanism that was described for the C4C4 RING domain of CNOT4.

In conclusion, the solution structure of p44-(321–395) shows that its topology differs from that of other reported RING domains by a circular permutation of the extended secondary structure elements. Site-directed mutagenesis suggests that the tight binding to p34 is mediated by hydrophobic interactions and involves residues on the solvent-exposed face of the  $\alpha$ -helix. The structure of p44-(321–395) sheds new light on the versatile nature of protein-protein interactions mediated by RING domains.

**Acknowledgments**—We express our gratitude to J. C. Thierry and J. M. Egly for constant support and fruitful discussions. We are grateful to C. Ling for the management of the computing and NMR facilities and to R. A. Atkinson for critical reading of the manuscript.

#### REFERENCES

- Frit, P., Bergmann, E., and Egly, J. M. (1999) *Biochimie (Paris)* **81**, 27–38
- Zurita, M., and Merino, C. (2003) *Trends Genet.* **19**, 578–584
- Egly, J. M. (2001) *FEBS Lett.* **24884**, 124–128
- Lehmann, A. R. (2003) *Biochimie (Paris)* **85**, 1101–1111
- Giglia-Mari, G., Coin, F., Ranish, J. A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N. G., Raams, A., Argentini, M., van der Spek, P. J., Botta, E., Stefanini, M., Egly, J. M., Aebersold, R., Hoeijmakers, J. H., and Vermeulen, W. (2004) *Nat. Genet.* **36**, 714–719
- Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J. M. (2000) *Cell* **102**, 599–607
- Chang, W. H., and Kornberg, R. D. (2000) *Cell* **102**, 609–613
- Andersen, G., Busso, D., Poterszman, A., Hwang, J. R., Wurtz, J. M., Ripp, R., Thierry, J. C., Egly, J. M., Moras, D., Svejstrup, J. Q., and Vichi, P. (1997) *EMBO J.* **16**, 958–967
- Gervais, V., Busso, D., Wasielewski, E., Poterszman, A., Egly, J. M., Thierry, J. C., and Kieffer, B. (2001) *J. Biol. Chem.* **276**, 7457–7464
- Gervais, V., Lamour, V., Jawhari, A., Frindel, F., Wasielewski, E., Dubaele, S., Egly, J. M., Thierry, J. C., Kieffer, B., and Poterszman, A. (2004) *Nat. Struct. Mol. Biol.* **11**, 616–622
- Seröz, T., Perez, C., Bergmann, E., Bradsher, J., and Egly, J. M. (2000) *J. Biol. Chem.* **275**, 33260–33266
- Tremeau-Bravard, A., Perez, C., and Egly, J. M. (2001) *J. Biol. Chem.* **276**, 27693–27697
- Fribourg, S., Romier, C., Werten, S., Gangloff, Y. G., Poterszman, A., and Moras, D. (2001) *J. Mol. Biol.* **306**, 363–373
- Fribourg, S., Kellenberger, E., Rogniaux, H., Poterszman, A., Van Dorsselaer, A., Thierry, J. C., Egly, J. M., Moras, D., and Kieffer, B. (2000) *J. Biol. Chem.* **275**, 31963–31971
- Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) *Protein Sci.* **2**, 543–558
- Vuister, G. W., Delaglio, F., and Bax, A. (1993) *J. Biomol. NMR* **3**, 67–80
- Bartels, C., Xia, T. H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR* **5**, 1–10
- Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54**, 905–921
- Linge, J. P., O'Donoghue, S. I., and Nilges, M. (2001) *Methods Enzymol.* **339**, 71–90
- Kuszewski, J., Gronenborn, A. M., and Clore, G. M. (1996) *Protein Sci.* **5**, 1067–1080
- Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160**, 65–73
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8**, 477–486
- Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graph.* **14**, 51–55
- Vriend, G. (1990) *J. Mol. Graph.* **8**, 52–56
- Crowley, P., Ubbink, M., and Otting, G. (2000) *J. Am. Chem. Soc.* **122**, 2968–2969
- Hanzawa, H., de Ruwe, M. J., Albert, T. K., van Der Vliet, P. C., Timmers, H. T., and Boelens, R. (2001) *J. Biol. Chem.* **276**, 10185–10190
- Summers, M. F., South, T. L., Kim, B., and Hare, D. R. (1990) *Biochemistry* **29**, 329–340
- Coleman, J. E. (1993) *Methods Enzymol.* **227**, 16–43
- Holm, L., and Sander, C. (1997) *Nucleic Acids Res.* **25**, 231–234
- Aravind, L., and Koonin, E. V. (2000) *Curr. Biol.* **10**, 132–134
- Ohi, M. D., Vander Kooi, C. W., Rosenberg, J. A., Chazin, W. J., and Gould, K. L. (2003) *Nat. Struct. Mol. Biol.* **10**, 250–255
- Bernstein, B. E., Hoffman, R. C., Horvath, S., Herriott, J. R., and Klevit, R. E. (1994) *Biochemistry* **33**, 4460–4470

33. Donate, L. E., Rufino, S. D., Canard, L. H., and Blundell, T. L. (1996) *Protein Sci.* **5**, 2600–2616
34. Lorick, K. L., Jensen, J. P., Fang, S., Ong, A. M., Hatakeyama, S., and Weissman, A. M. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 11364–11369
35. VanDemark, A. P., and Hill, C. P. (2002) *Curr. Opin. Struct. Biol.* **12**, 822–830
36. Dominguez, C., Folkers, G. E., and Boelens, R. (2004) *Handbook of Metalloproteins*, Vol. 3, pp. 338–351, Wiley & Sons, New York
37. Albert, T. K., Hanzawa, H., Legtenberg, Y. I., de Ruwe, M. J., van den Heuvel, F. A., Collart, M. A., Boelens, R., and Timmers, H. T. (2002) *EMBO J.* **21**, 355–364
38. Dominguez, C., Bonvin, A. M., Winkler, G. S., van Schaik, F. M., Timmers, H. T., and Boelens, R. (2004) *Structure* **12**, 633–644
39. Winkler, G. S., Albert, T. K., Dominguez, C., Legtenberg, Y. I., Boelens, R., and Timmers, H. T. (2004) *J. Mol. Biol.* **337**, 157–165
40. Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B., and Friedberg, E. C. (1996) *Biochemistry* **35**, 2157–2167
41. Coin, F., Marinoni, J. C., Rodolfo, C., Fribourg, S., Pedrini, A. M., and Egly, J. M. (1998) *Nat. Genet.* **20**, 184–188
42. Zheng, N., Wang, P., Jeffrey, P. D., and Pavletich, N. P. (2000) *Cell* **102**, 533–539