CBP/p300 TAZ1 Domain Forms a Structured Scaffold for Ligand Binding[†,‡]

Roberto N. De Guzman, Jonathan M. Wojciak, Maria A. Martinez-Yamout, H. Jane Dyson, and Peter E. Wright*

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

Received August 25, 2004; Revised Manuscript Received October 8, 2004

ABSTRACT: The transcriptional coactivator protein CBP and its paralog p300 each contain two homologous zinc-containing TAZ domains, which constitute the interaction sites for a number of transcription factors. Previous reports of the three-dimensional structures of TAZ1 in complex with binding partners and of the isolated CBP TAZ2 domain show a distinctive topology composed of four amphipathic helices, organized by three zinc-binding clusters with HCCC-type coordination. The isolated CBP TAZ2 domain forms a stable three-dimensional structure in solution, but a recent report [Dial, R., Sun, Z., and Freedman, S. J. (2003) *Biochemistry 42,* 9937] suggested that the isolated p300 TAZ1 domain lacks a well-defined structure and behaves like a molten globule, even in the presence of Zn^{2+}, and that the formation of a stable three-dimensional structure requires binding of a protein partner. In marked contrast to this result, we find that both the CBP and p300 TAZ domains in the presence of stoichiometric concentrations of Zn^{2+} adopt a well-defined structure in solution in the absence of binding partners. We have determined the three-dimensional structure of the isolated CBP TAZ1 domain by NMR methods and show that it has the same structure in the presence and absence of binding partners. This is an important finding: whether the free TAZ1 domain forms a folded structure or behaves as a molten globule will have a significant bearing on the mechanism of protein−protein recognition. Although TAZ1 and TAZ2 share many structural similarities, there is a major structural difference: the fourth helix is oriented in opposite directions in the TAZ1 and TAZ2 domains. The structure of the free TAZ1 domain suggests that this difference is an inherent feature that determines binding specificity and facilitates discrimination between different subsets of transcription factors by the two TAZ domains.

The transcriptional coactivator proteins CBP [cyclic-AMP response element binding protein (CREB) binding protein][1] (*1*) and its paralog p300 (*2*) are relatively large proteins (2441 residues in human CBP) and contain several protein-binding domains, as well as an acetyltransferase enzyme activity. CBP and p300 contain two homologous TAZ (transcriptional adaptor zinc-binding) (*3*) domains, TAZ1 and TAZ2, which are important sites of protein−protein interaction. TAZ1 and TAZ2 each contain three highly homologous zinc-binding sites with the consensus sequence His1-X_3-Cys2-X_{4-12}-Cys3-X_{2-4}-Cys4 (where X is any amino acid and the subscript denotes the number of residues between the zinc ligands). The primary function of the TAZ domains appears to be in protein recognition, and currently, more than 30 different transcription factors have been reported to associate with CBP/p300 via its TAZ domains. Although the TAZ1 and TAZ2 domains share sequence homology, they bind different subsets of transcription factors.

The NMR structure of the CBP TAZ2 domain revealed a novel structure for a zinc-binding motif (*4*), composed of four amphipathic α-helices packed against each other to form a hydrophobic core. In the absence of zinc, the TAZ2 domain is unfolded, and binding of 3 equiv of zinc is required to fold the protein into a stable three-dimensional structure (*4*). The solution structures of TAZ1 bound to the activation domains of HIF-1α (hypoxia inducible factor) (*5*, *6*) and CITED2 (CBP/p300-interacting transactivator with ED-rich tail) (*7*, *8*) indicate that TAZ1 adopts a three-dimensional structure with a similar fold to that of the TAZ2 domain. The free activation domains of both HIF-1α and CITED2 are intrinsically unstructured and fold only upon binding to TAZ1.

In contrast to the TAZ2 domain, which is structured in the absence of ligands (*4*), the free TAZ1 domain of p300 has been reported to be only partially folded even in the presence of saturating amounts of zinc, behaving more like a molten globule than a globular protein (*9*). On the basis of the poor dispersion of the NMR $^1H-^{15}N$ correlation spectrum

[†] This work was supported by grant CA96865 from the National Institutes of Health to P.E.W. R.N.D. is supported by a postdoctoral fellowship from the Leukemia and Lymphoma Society. J.M.W is supported by a postdoctoral fellowship from the American Cancer Society.

[‡] The coordinates have been deposited in the Protein Data Bank with accession number 1U2N and the chemical shift assignments in the BioMagResBank with accession number 6268.

* Corresponding author. Department of Molecular Biology MB2, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037. Phone: 858-784 9721. Fax: 858-784 9822. E-mail: wright@scripps.edu.

[1] Abbreviations: CBP, CREB-binding protein; CD, circular dichroism; CITED, CBP/p300-interacting transactivator with ED-rich tail; CREB, cyclic AMP-response element binding protein; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; HIF, hypoxia inducible factor; HSQC, heteronuclear single-quantum coherence; IPTG, isopropyl-β-D-thiogalactopyranoside; MES, 2-[*N*-morpholino]ethanesulfonic acid; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; TAZ, transcriptional adaptor zinc-binding; τ_m, mixing time.

10.1021/bi048161t CCC: $30.25
Published on Web 12/22/2004

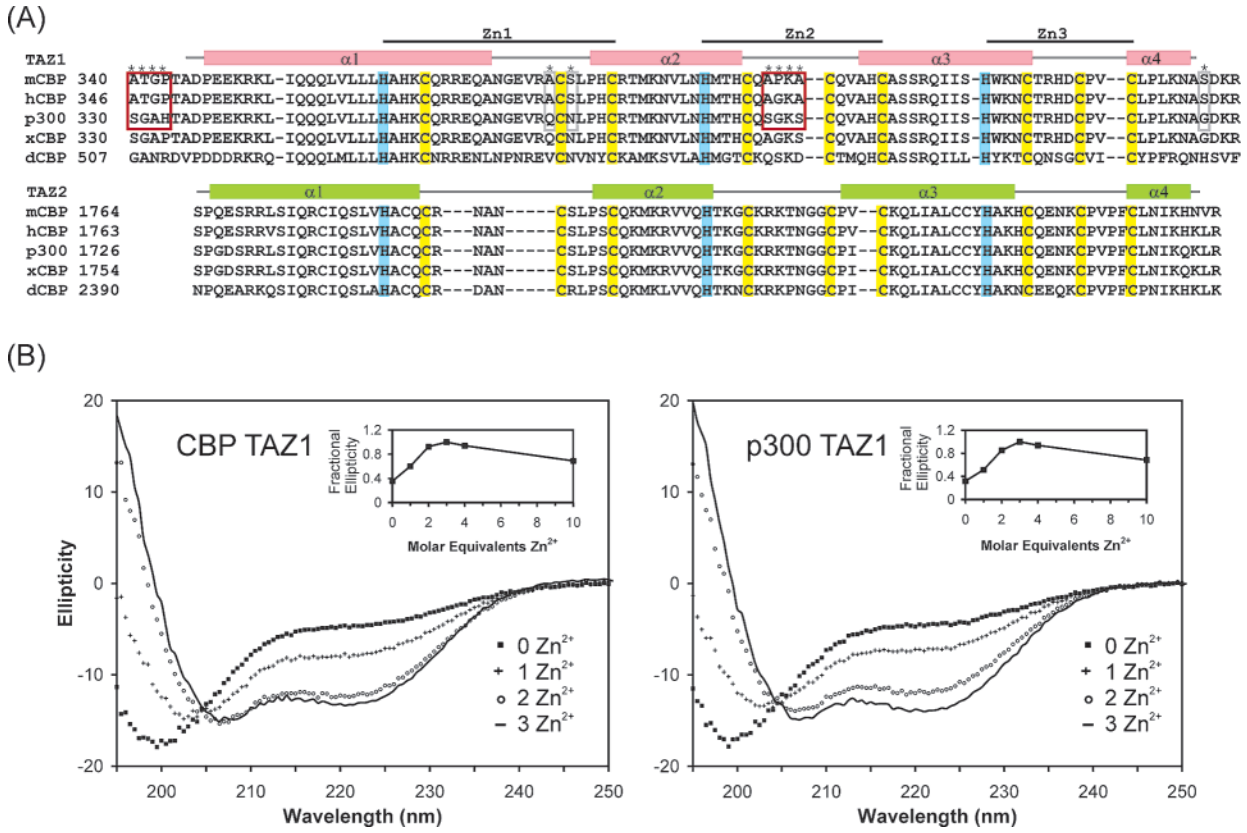

FIGURE 1: (A) Protein sequence alignment of the CBP and human p300 TAZ1 and TAZ2 domains from *Mus musculus* (m), *Homo sapiens* (h), *Xenopus laevis* (x), and *Drosophila melanogaster* (d). The zinc-coordinating histidine and cysteine residues are highlighted, and the three zinc-binding (Zn1−Zn3) clusters are marked by solid lines. The secondary structural elements of TAZ1 and TAZ2 are shown as solid bars for the four helices ($\alpha_1-\alpha_4$) and solid lines for the loops. The differences in the mouse, human, and p300 TAZ1 amino acid sequence are boxed and shown with asterisks. (B) Far-UV circular dichroism spectra of CBP (left panel) and p300 (right panel) TAZ1 with increasing equivalents of Zn^{2+}. The spectrum of zinc-free TAZ1 is indicated by squares; the spectra upon addition of 1, 2 and 3 equiv of Zn^{2+} are shown by crosses, circles, and the solid line, respectively. (Insets) Plot of mean residue ellipticity (MRE) at 222 nm as a percentage of the difference between the value in the absence of zinc (set at 0%) and the value in the presence of 3 molar equiv of zinc (set at 100%).

(HSQC) of p300 TAZ1 in the presence of Zn^{2+}, it was suggested that TAZ1 requires binding to either HIF-1α or CITED2 to fold into a stable three-dimensional structure (*9*). These results are inconsistent with experimental data obtained in this laboratory for the highly homologous TAZ1 domains from CBP and p300 (Figure 1A). We report here that the isolated TAZ1 domains from both CBP and p300 fold into stable globular structures in the presence of Zn^{2+} ions. Furthermore, we also describe the solution structure of the isolated CBP TAZ1 domain, which provides insights into the mechanism by which the TAZ1 and TAZ2 domains discriminate between different transcriptional activation domains. Our results show unequivocally that protein−protein interactions are not required for the CBP TAZ1 domain to fold.

MATERIALS AND METHODS

Sample Preparation. TAZ1 domains from mouse CBP (residues 340−439, 100% identity with human sequence) and human p300 (residues 323−424) were expressed in three different ways. The method that gives the greatest yield of protein involves coexpression with the activation domain of HIF-1α or CITED2, purification of the complex from the soluble fraction of the cell lysate, followed by purification of the isolated TAZ1 by reversed phase HPLC, and refolding in the presence of Zn^{2+} as previously described (*5*, *8*). In the second method, CBP and p300 TAZ1 were expressed individually using the expression vector pET21a (Novagen) and purified from inclusion bodies as previously described for TAZ2 (*4*). In the third method, folded CBP and p300 TAZ1 were isolated under native conditions from both individual and coexpression systems, using a low-temperature expression protocol in which *E. coli* BL21(DE3) (DNAY) were grown in LB or M9 minimal medium supplemented with 150 μM $ZnSO_4$ at 37 °C to OD_{600} = 0.7. Protein expression was induced by addition of 1 mM IPTG and growth was continued at 15 °C for an additional 12−16 h. Cells were lysed by sonication, and soluble protein was isolated by ion exchange chromatography (HiTrap SP, Amersham Biosciences) in 20 mM Tris pH7.5, 50 mM NaCl, 10 μM $ZnSO_4$, 2 mM DTT using a linear gradient to 600 mM NaCl. For NMR analysis, native TAZ1 was dialyzed into 10 mM Tris pH 6.9, 50 mM NaCl, 2 mM DTT. The TAZ1 domain expressed using the three different protocols behaved identically under all circumstances.

Circular Dichroism Spectroscopy. CD spectra were collected using an Aviv model 202 CD spectrometer at 25 °C, using a 0.2 cm cell. Zinc titration was performed using CBP and p300 TAZ1 purified under denaturing conditions to remove the bound Zn^{2+}. The lyophilized TAZ1 was dissolved in 3 mM Tris-HCl pH 7.5 to give a protein concentration of 5−10 μM, and increasing equivalents of Zn^{2+} were added.

NMR Spectroscopy. NMR spectra were acquired at 25 °C on Bruker DRX600, DRX800, and AVANCE 900 MHz

Table 1: Experimental Restraints and Structural Statistics for the CBP TAZ1 Domain

NMR Restraints	
total distance restraints	1070
intraresidue	386
sequential	317
medium range	250
long range	117
total dihedral angle restraints	179
φ	84
ψ	63
χ_1	32
Ensemble Statistics (20 structures)	
violation analysis	
maximum distance violation	0.35 Å
maximum dihedral angle violation	0.70°
energies	
mean restraint energy	36 kcal mol^{-1}
mean AMBER energy	−4157 kcal mol^{-1}
rms deviation from the mean structure (helices only)	
backbone atoms (N,Cα,C)	0.48 Å
all heavy atoms	0.98 Å
deviation from idealized geometry	
bond lengths	0.01 Å
bond angles	3.2°
Ramachandran plot	
most favorable regions	83.3%
additionally allowed regions	15.1%
generously allowed regions	1.4%
disallowed regions	0.2%

spectrometers, processed using NMRPipe (*10*), and analyzed using NMRView (*11*). Resonance assignments for the isolated CBP TAZ1 domain were obtained from heteronuclear (^{1}H,^{15}N,^{13}C) and multidimensional NMR experiments as previously described for the TAZ1/HIF-1α (*5*) and TAZ1/CITED2 (*8*) complexes. Distance restraints were obtained from 3D ^{15}N-edited NOESY−HSQC (τ_m 120 ms) and 3D ^{13}C-edited NOESY−HSQC (τ_m 120 ms) acquired for samples in 100% ^2H_2O and 10% ^2H_2O.

Structure Calculations and Analysis. Interproton upper bound distance restraints were derived from cross-peak volumes in the NOESY experiments and were assigned upper bounds of 2.7, 3.5, 4.5, and 5.5 Å, with 1.8 Å lower bounds. Torsion angle restraints for φ, ψ, and χ_1 were obtained from the chemical shift index and intraresidue NOE connectivities (*12*−*14*). Distance restraints were obtained by manual analysis of NOESY spectra, and additional distance restraints were generated with the program SANE (*15*) and added during the structure calculation. The Zn^{2+}-bound cysteine and histidine residues were identified on the basis of the chemical shifts and sequence homology of the zinc-binding clusters of TAZ1 and TAZ2 domains (Figure 1A). All nine cysteine residues in CBP TAZ1 have downfield $^{13}C^{\beta}$ chemical shifts in the range of 29.1−31.0 ppm, which is indicative of Zn^{2+}-coordination (*5*). Tetrahedral geometry was imposed for the Zn^{2+}-ligands during the structure calculation by distance and torsion restraints. Hydrogen-bond restraints were used during the initial model building for regions that were identified by the chemical shift index and NOE connectivities to be helices. The total number of distance and torsion angle restraints used is summarized in Table 1. An initial set of 100 simulated annealing structures was generated using DYANA (*16*) followed by variable target function minimization. The 100 DYANA structures with the lowest target function were further refined by restrained molecular dynamics in AMBER7 using the *ff99* force field (*17*). When the distance and dihedral angle violations dropped below 100 kcal/mol, refinement was performed using the generalized Born (GB) potential in AMBER7 to account for the effects of solvent (*18*). Twenty structures with the lowest combined distance and angle violations were sorted by AMBER energy and selected for analysis. Backbone analysis was performed using PROCHECK (*19*), and graphics images were prepared using MOLMOL (*20*).

RESULTS

Folding of CBP/p300 with Zn^{2+}. Two of the three protocols used to prepare CBP or p300 TAZ1 involved purification by reversed phase HPLC, resulting in an unfolded, zinc-free protein. The third protocol led directly to folded, zinc-containing protein, which loses zinc and unfolds upon treatment with excess EDTA. The addition of 3 equiv of Zn^{2+} to the unfolded protein prepared by any of these three methods led to identical samples of TAZ1 that are, by CD and NMR spectroscopic criteria, stably folded. The effect of the addition of Zn^{2+} to CBP and p300 TAZ1 is shown in Figure 1B. The CD spectrum of zinc-free TAZ1 is characteristic of a random coil state lacking secondary structural elements. Addition of up to 3 equiv of Zn^{2+} results in an increase in the molar ellipticity at 222 and 208 nm and a decrease in the random-coil signal below 200 nm, consistent with an increase in the α-helical content of TAZ1. The CD spectra show maximal helicity at 3 equiv of Zn^{2+} and decreased helicity at higher zinc concentrations (Figure 1B, insets), suggesting that excess Zn^{2+} results in misfolding of TAZ1, probably due to nonspecific metal binding. The presence of an approximate isodichroic point in the CD spectra of Figure 1B suggests that zinc binding is close to being a two-state process and that stable intermediates with fewer than 3 equiv of zinc are not formed in high concentration.

Titration of the CBP TAZ1 domain by Zn^{2+} was also monitored by 2D NMR spectroscopy using ^{15}N-labeled TAZ1 (Figure 2). In the absence of Zn^{2+}, the narrow range of amide proton chemical shifts in the 2D ^{1}H−^{15}N HSQC spectrum shows that TAZ1 is unstructured in solution (Figure 2A). Upon addition of 3 equiv of Zn^{2+}, the line width and proton resonance dispersion in the HSQC spectrum are indicative of a folded three-dimensional structure (Figure 2C). In the presence of substoichiometric amounts of Zn^{2+}, the HSQC spectrum reveals the presence of both unfolded and fully folded TAZ1 (Figure 2B). There also appear to be additional low-intensity peaks in Figure 2B, which may be indicative of intermediate structures partly populated by zinc (see, for example, the tryptophan side chain cross-peak marked with an arrow in Figure 2B). With saturating amounts of Zn^{2+}, the intermediate peaks and peaks corresponding to the unfolded TAZ1 disappear and peaks corresponding to the folded TAZ1 domain become the predominant feature of the HSQC (Figure 2C). When bacterially expressed TAZ1 is purified under native conditions in the presence of Zn^{2+}, the 2D HSQC spectrum is identical to that of the refolded TAZ1 domain after the addition of 3 equiv of zinc (data not shown), indicating that the refolding procedure has no effect on the structure.

The amino acid sequences of the CBP and p300 TAZ1 domains are almost identical, with only a few conservative

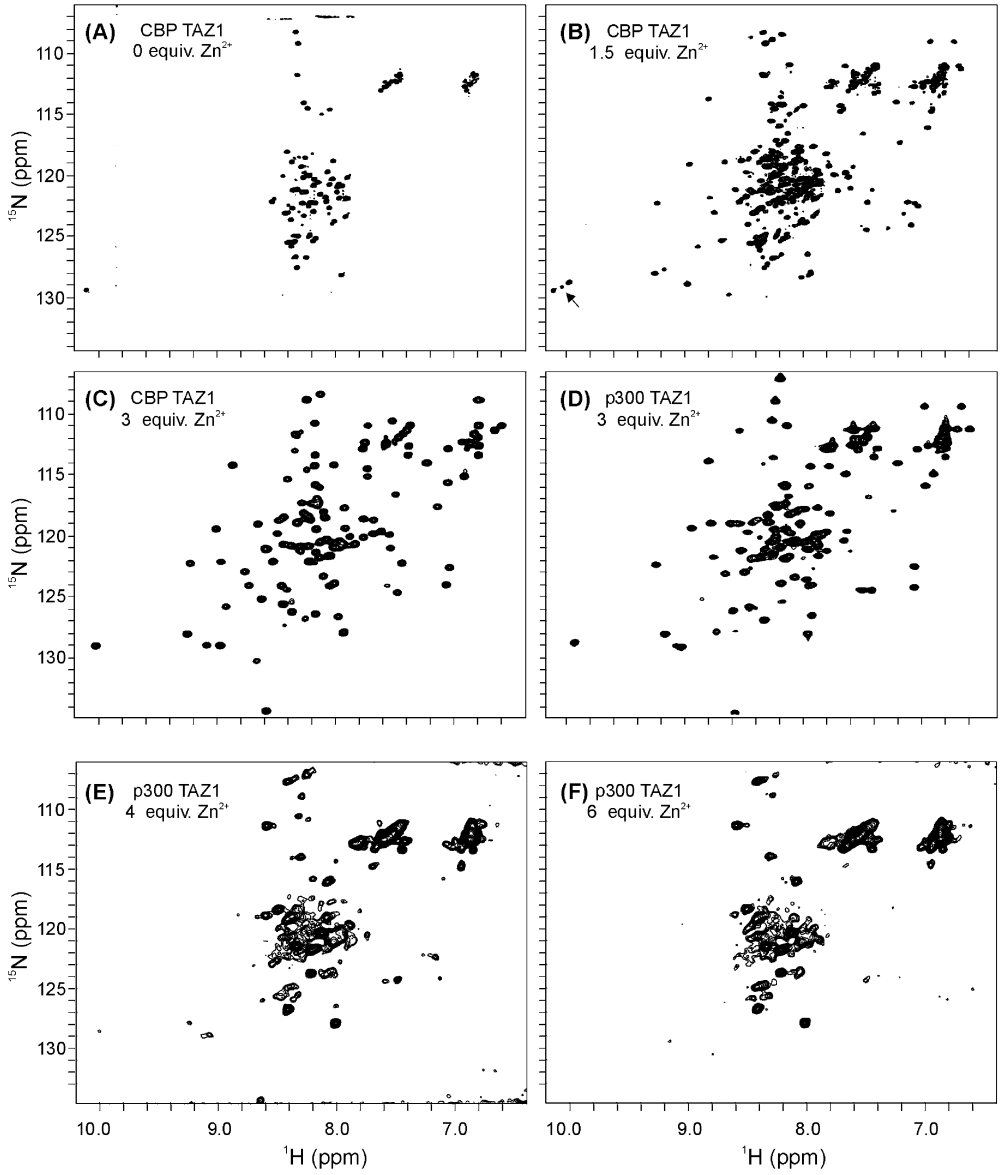

FIGURE 2: 600 MHz $^{15}N-^{1}H$ HSQC spectra of TAZ1 domains. (A) CBP TAZ1 (200 μM) purified under denaturing conditions (no zinc bound), in 10 mM Tris-d_{11}, 5 mM DTT-d_{10}, pH 6.8. (B) CBP TAZ1 (200 μM in 10 mM Tris-d_{11}, 5 mM DTT-d_{10}, pH 6.8) after addition of 1.5 equiv of Zn^{2+} (final concentration 300 μM $ZnSO_4$,) (C) CBP TAZ1 (1 mM CBP TAZ1 in 10 mM Tris-d_{11}, 5 mM DTT-d_{10}, pH 6.3) after titration of 3 equiv of Zn^{2+} (final concentration 3 mM $ZnSO_4$). (D) p300 TAZ1 purified under native conditions (0.45 mM p300 TAZ1 containing zinc, 10 mM Tris, 50 mM NaCl, 1 mM DTT, pH 6.9). (E, F) p300 ^{15}N-TAZ1 in 5 mM MES, 100 mM NaCl, pH 6 (as used by Dial et al. (*9*)), in the presence of (E) 4 equiv and (F) 6 equiv of Zn^{2+}.

changes between the two sequences (Figure 1A). It was therefore a surprise that the p300 TAZ1 domain was reported to be unstructured, even in the presence of Zn^{2+}, and required binding of the HIF-1α activation domain to fold into a stable three-dimensional structure (*9*). In our hands, the recombinant p300 ^{15}N-TAZ1 (residues 323−424) purified under native conditions showed a highly dispersed HSQC spectrum diagnostic of a folded globular protein domain (Figure 2D). The spectrum is very similar to that of the CBP TAZ1 domain; the observed chemical shift differences can be attributed to the eight amino acid substitutions between p300 and CBP TAZ1 (Figure 1A) and definitely do not reflect major differences in structure. Thus, the HSQC spectra indicate that the free p300 TAZ1 domain folds into a well-defined three-dimensional structure and does not behave like a molten globule as previously reported (*9*). The overall similarity of the HSQC spectra of the p300 and CBP TAZ1 domains suggests that their structures are very similar. Like CBP, the p300 TAZ1 domain can also be purified under denaturing conditions and refolded in Tris buffer in the presence of DTT and 3 equiv of Zn^{2+} to give an HSQC spectrum (not shown) that is identical to that of Figure 2D. A very similar, well-dispersed HSQC spectrum is obtained when the p300 TAZ1 domain folded with stoichiometric amounts of Zn^{2+} is exchanged into the buffer used by Dial et al. (*9*) (5 mM MES/100 mM NaCl, pH 6.0). However, addition of excess Zn^{2+} under these buffer conditions leads to loss of dispersion and line broadening and results in spectra (Figure 2E,F) very similar to those reported (*9*).

Solution Structure of CBP TAZ1. The solution structure of the free CBP TAZ1 domain determined from the NMR constraints is shown in Figure 3. The family of structures is well-defined, especially in the helical regions, with some disorder in the loop between α_1 and α_2. The degree of

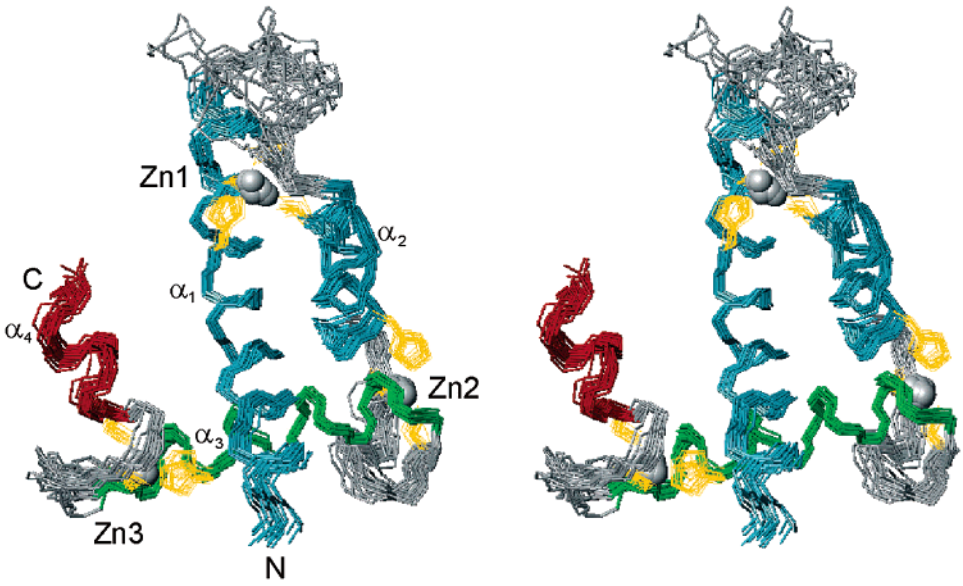

FIGURE 3: Stereoview showing the best-fit superposition of backbone heavy atoms in the 20 lowest energy NMR structures of the isolated CBP TAZ1 domain. Helices α_1 and α_2 are shown in blue, helix α_3 in green, and helix α_4 in red. The side chains of the ligand residues in each zinc binding site (Zn1−Zn3) are shown in yellow, and the zinc atoms are shown as pale gray spheres. The amino (N) and carboxyl (C) termini are indicated. The figure was prepared using MOLMOL (*20*).

constraint satisfaction and the Ramachandran statistics are summarized in Table 1. The overall structure is very similar to that of the TAZ1 domain bound to the activation domains of HIF-1α (*5*, *6*) and CITED2 (*7*, *8*). The structure of TAZ1 contains four amphipathic helices ($\alpha_1-\alpha_4$) organized by three zinc binding motifs (Zn1−Zn3). Each zinc binding motif consists of two helices joined by a connecting loop, with one histidine and three cysteine ligands that are arranged sequentially to form an HCCC-type zinc-binding motif, similar to the zinc-binding motifs found in TAZ2 (*4*). In each of the zinc-binding motifs, the first two zinc ligands (His and Cys) are located at the C-terminus of the first helix, the third ligand (Cys) is on the connecting loop, and the fourth ligand (Cys) is located at the N-terminus of the second helix. The three zinc-binding clusters (Zn1−Zn3) are distributed approximately at the apexes of an irregular triangle with the four amphipathic helices packed through hydrophobic interactions to form the core of the domain.

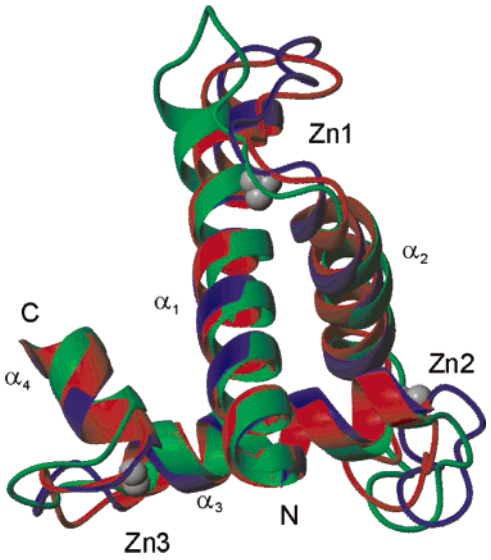

FIGURE 4: Superposition of the ribbon structures of the CBP TAZ1 domain when free (red), complexed with HIF-1α (green), and complexed with CITED2 (blue). The lowest energy structure from the NMR ensemble is used in each case. The figure was prepared using MOLMOL (*20*).

DISCUSSION

Why Study the Free TAZ1 Domain? As the binding site for at least a dozen different transcription factors, the CBP/p300 TAZ1 domain plays an important role in protein−protein interactions involved in the transcriptional regulation of many cellular pathways. There are NMR structures of the CBP/p300 TAZ1 domain in complex with the activation domains of HIF-1α (*5*, *6*) and CITED2 (*7*, *8*). When bound to HIF-1α or CITED2, the TAZ1 domain adopts a three-dimensional structure similar to that of the homologous TAZ2 domain (*4*). Because of the similarities of the structure of TAZ1 in the HIF-1α and CITED2 complexes, together with the structural similarities between TAZ1 and the free TAZ2, it was generally assumed that the structure of TAZ1 in the free and bound forms would be the same and that there would be no major structural rearrangement in the TAZ1 structure induced by protein binding.

We report here that the free CBP/p300 TAZ1 domain adopts a fully folded three-dimensional structure, which is not in agreement with the report of Dial et al. (*9*) that the free p300 TAZ1 domain is a molten globule. It is important to know whether the free TAZ1 domain forms a folded structure or behaves as a molten globule, because this will have a significant bearing on the mechanism of protein−protein recognition by the TAZ domains of CBP and p300. An accurate picture of the folded state of free TAZ1 is important for distinguishing between two different but equally plausible mechanisms of protein−protein recognition. First, if TAZ1 is not well structured (molten globule or unfolded) in isolation, then a mechanism of mutual synergistic folding could be invoked. This mechanism appears to operate in the binding of the p160 domain ACTR to another CBP domain, the nuclear coactivator binding domain (NCBD) (*21*), and was invoked by Dial et al. (*9*) for the interaction

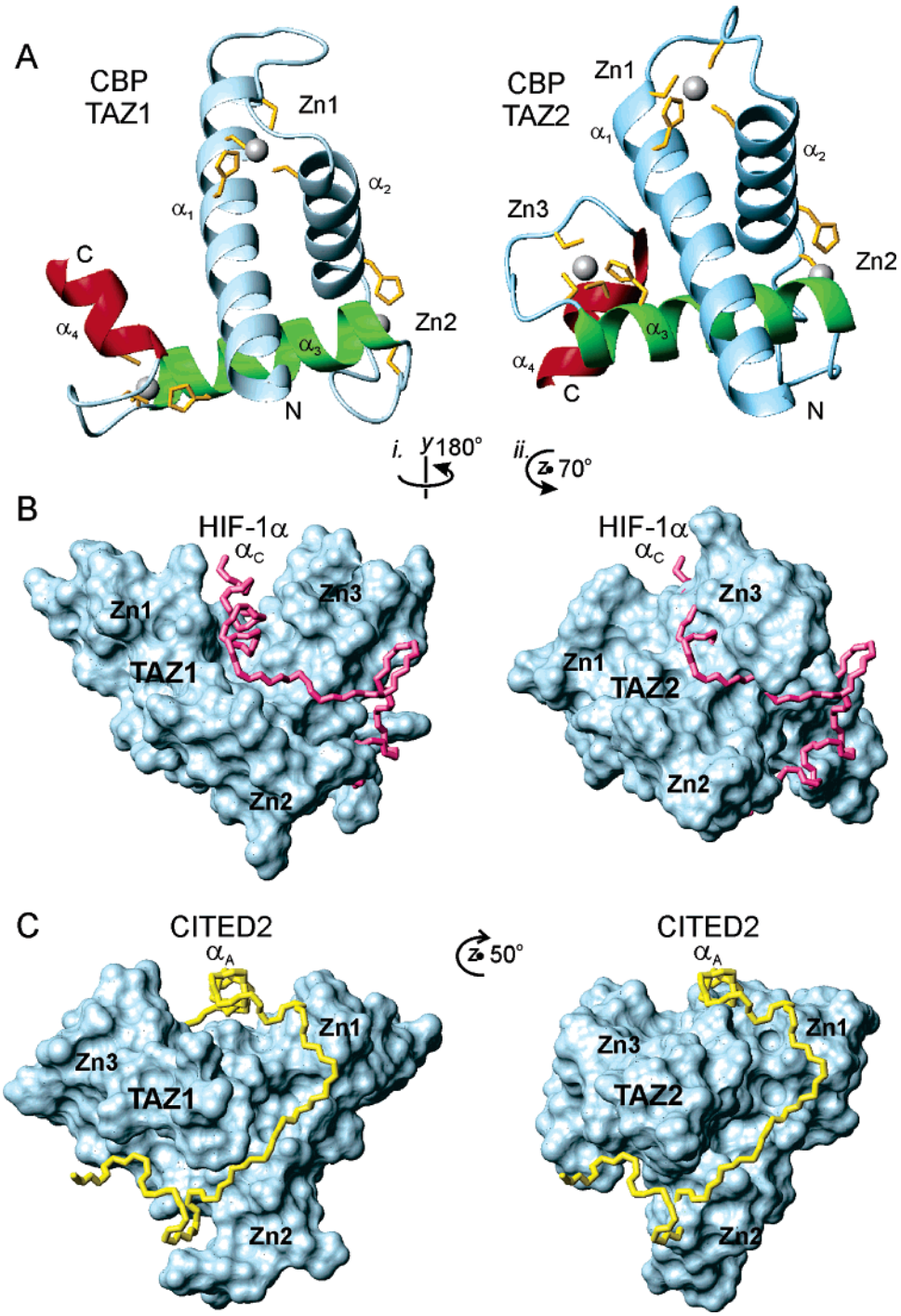

FIGURE 5: (A) Ribbon diagrams showing the backbone structures of free TAZ1 (left) and free TAZ2 (right) domains of CBP. (Color code as in Figure 3.) (B) (left) Surface representation of TAZ1 in complex with HIF-1α (pink), showing the deep groove that binds the α_C helix of HIF-1α (from the structure PDB 1L8C (*5*)). (right) Surface representation of TAZ2 with HIF-1α (pink) modeled in the analogous position, showing occlusion of the binding groove for the α_C helix. The coordinates in panel A were rotated by 180° along the *y*-axis followed by a 70° rotation along the *z* axis to obtain the orientation shown in panel B. (C) (left) Surface representation of TAZ1 in complex with CITED2 (yellow), showing the well-defined groove that accommodates the α_A helix of CITED2 (from the structure PDB 1R8U (*8*)). (right) Surface representation of TAZ2, with CITED2 (yellow) modeled in the analogous position: the absence of the helix α_A binding groove at the top of the structure would disfavor CITED2 binding. The coordinates in panel A were rotated by 50° along the *z*-axis to obtain the orientation shown in panel C. The figure was prepared using MOLMOL (*20*).

of TAZ1 with ligands. The second alternative involves templated folding of unstructured ligands on a preformed scaffold. This mechanism appears to operate with the interaction of the phosphorylated kinase-inducible domain (pKID) of CREB with another CBP domain, the KIX domain (*22*). Our results clearly distinguish between these two possibilities for TAZ1.

Finally, whether TAZ1 is fully structured or a molten globule is directly relevant to our understanding of the overall structure of CBP and p300: a structured TAZ1 domain would mean that CBP is composed of several structured functional domains flanked by disordered regions, much like beads on a string. On the other hand, if the TAZ1 domain is a molten globule, then the entire N-terminal region of CBP,

nearly 600 amino acids, would be disordered and behave as a long flexible string that forms compact domains only in the presence of binding partners.

The Free TAZ1 Domains of CBP and p300 Are Structured. The experiments reported in the present paper show unequivocally that the free TAZ1 domain, from either p300 or CBP, whether expressed under native conditions or refolded by addition of Zn^{2+} to the denatured protein, adopts a well-defined structure and does not require interactions with target proteins to form a stable fold. Both circular dichroism and NMR spectra show that, in the absence of Zn^{2+}, the TAZ1 domains of CBP and p300 are unstructured and that binding of 3 equiv of Zn^{2+} is required for folding. The NMR and CD spectra also show that zinc-induced folding is approximately two-state and that the concentration of stable intermediate forms at substoichiometric (i.e., at less than 3 equiv) concentrations of Zn^{2+} is low.

Structure of TAZ1. The structure of the free CBP TAZ1 domain closely resembles that of TAZ1 complexed with the activation domains of HIF-1α and CITED2 (Figure 4). The root-mean-square deviations for backbone heavy atoms in the helical regions of the free TAZ1 domain compared to complexes with HIF-1α and CITED2 are 1.14 ± 0.17 and 1.21 ± 0.17 Å, respectively (for ensembles of 20 structures each). This shows that no significant structural rearrangement occurs upon binding to target proteins. Thus, TAZ1 acts as a fully structured template for binding the intrinsically disordered activation domains of HIF-1α and CITED. The structural differences in the interhelical loops apparent in Figure 4 are probably not significant, since these regions are somewhat disordered in the solution structures and show signs of enhanced dynamics in relaxation experiments (J. Lansing and P. E. Wright, unpublished data).

Comparison of TAZ1 and TAZ2 Structures. Both TAZ domains are comprised of four amphipathic helices that pack to form a hydrophobic core (Figure 5A). A HCCC zinc coordination site is located at the junction between each pair of helices and binding of zinc is required for stabilization of the folded structure—both the TAZ1 and TAZ2 motifs are unstructured in the absence of zinc. The first three helices, $\alpha_1-\alpha_3$, are packed in the same topological arrangement in both domains. Although the TAZ1 α_1 helix is longer by five residues than the corresponding helix in TAZ2 α_1, the relative position of the zinc atom in both domains is very similar. The major structural difference between TAZ1 and TAZ2 is in the different orientation of the fourth helix, α_4 (red in Figure 5A). In TAZ1, the α_1 and α_4 helices are both located on the same side of helix α_3, while in TAZ2 these helices are on opposite sides of α_3. The difference in the orientation of the fourth helix also results in a major shift in the position of the zinc atom in Zn3, relative to Zn1 and Zn2. This structural difference is also reflected in the number of interhelical contacts: there are more interhelical NOEs between α_1 and α_4 in TAZ1 than in TAZ2. There are 27 long-range NOEs between helix α_1 residues Gln354, Leu357, Val358, and Leu361, and helix α_4 residues Val428 and Leu432. In TAZ2, the arrangement of the four helices results in a structure approximating a pyramid in overall shape, whereas the overall shape of TAZ1 is more flattened and triangular due to the positioning of α_4. The interactions that determine the orientation of helix α_4 are not yet known. It has been hypothesized that differences in spacing between the last two cysteine ligands of Zn3 may influence the orientation of α_4 (*5*); in TAZ1, the ligands are separated by two residues (PV), compared to four residues (PVPF) in TAZ2 (Figure 1A). However, other explanations are possible, and we are currently conducting mutagenesis experiments to address this issue.

Since the orientation of helix α_4 of TAZ1 is similar to that seen in the HIF-1α and CITED2 complexes, it appears that the structural differences between TAZ1 and TAZ2 are an inherent feature that contributes to their binding specificity and their ability to recognize different subsets of transcription factors. The altered packing of the helices results in significant differences in the characteristic grooves on the protein surface that function as binding sites for CBP/p300 ligands (Figure 5B,C). For example, the deep groove in TAZ1 that accommodates the α_C helix of HIF-1α (*5*, *6*) is occluded by the α_4 helix of TAZ2 (Figure 5B), providing an explanation for the ability of the HIF-1α activation domain to discriminate between TAZ1 and TAZ2 in binding to CBP/p300. Similarly, the N-terminal helix of CITED2 binds in a deep groove formed by the α_1 and α_4 helices and the $\alpha_1-\alpha_2$ connecting loop in the TAZ1 structure (*7*, *8*); this groove is absent from the surface of TAZ2 because of the repositioning of the α_4 helix (Figure 5C). In addition to changes in surface shape, differences in the amino acid sequence lead to changes in the surface hydrophobicity that can influence the binding properties. Although the grooves that bind the LPXL motifs of HIF-1α and CITED2 (formed by α_1/α_2 and α_1/α_3 packing) are present in both TAZ1 and TAZ2, the hydrophobic residues that form the binding interface are replaced in TAZ2, thereby discriminating against LPXL binding.

Role of the TAZ1 Domain in CBP/p300 Protein Recognition. The data presented herein show that the TAZ1 domain is independently folded in the presence of zinc and does not require binding to target proteins to form a stable three-dimensional structure. Thus, the isolated TAZ1 domain behaves as a preformed scaffold, exhibiting surfaces and grooves that mediate protein−protein recognition. Previous studies have shown that the activation domains of HIF-1α and CITED2 are intrinsically unstructured and undergo coupled folding events on binding to TAZ1 (*5*−*8*). Our present structure shows that no significant conformational changes occur in TAZ1 upon binding to HIF-1α and CITED; the pronounced grooves in the surface of TAZ1 that are used for binding these targets are preformed in the uncomplexed protein.

ACKNOWLEDGMENT

We thank John Chung, Gerard Kroon, Brian Lee, and Sonja Dames for help with NMR experiments. We thank Melissa Allen, Cheyenne Reyes, Mindy Landes, and Linda Tennant for technical assistance.

REFERENCES

1. Chrivia, J. C., Kwok, R. P., Lamb, N., Hagiwara, M., Montminy, M., and Goodman, R. H. (1993) Phosphorylated CREB binds specifically to nuclear protein CBP, *Nature 365*, 855−859.
2. Eckner, R., Ewen, M. E., Newsome, D., Gerdes, M., DeCaprio, J. A., Lawrence, J. B., and Livingston, D. M. (1994) Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor, *Genes Devel. 8*, 869−884.

3. Ponting, C. P., Blake, D. J., Davies, K. E., Kendrick-Jones, J., and Winder, S. J. (1996) ZZ and TAZ: New putative zinc fingers in dystrophin and other proteins, *Trends Biochem. Sci. 21*, 11−13.
4. De Guzman, R. N., Liu, H. Y., Martinez-Yamout, M., Dyson, H. J., and Wright, P. E. (2000) Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP, *J. Mol. Biol. 303*, 243−253.
5. Dames, S. A., Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J., and Wright, P. E. (2002) Structural basis for Hif-1 alpha/CBP recognition in the cellular hypoxic response, *Proc. Natl. Acad. Sci. U.S.A 99*, 5271−5276.
6. Freedman, S. J., Sun, Z. Y., Poy, F., Kung, A. L., Livingston, D. M., Wagner, G., and Eck, M. J. (2002) Structural basis for recruitment of CBP/ p300 by hypoxia-inducible factor-1alpha, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5367−5372.
7. Freedman, S. J., Sun, Z. Y., Kung, A. L., France, D. S., Wagner, G., and Eck, M. J. (2003) Structural basis for negative regulation of hypoxia-inducible factor-1alpha by CITED2, *Nat. Struct. Biol. 10*, 504−512.
8. De Guzman, R. N., Martinez-Yamout, M., Dyson, H. J., and Wright, P. E. (2004) Interaction of the TAZ1 Domain of CREB−Binding Protein with the Activation Domain of CITED2: Regulation by Competition between Intrinsically Unstructured Ligands for Non-Identical Binding Sites, *J. Biol. Chem. 279*, 3042−3049.
9. Dial, R., Sun, Z. Y., and Freedman, S. J. (2003) Three conformational states of the p300 CH1 domain define its functional properties, *Biochemistry 42*, 9937−9945.
10. Delaglio, F., Grzesiek, S., Vuister, G. W., Guang, Z., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
11. Johnson, B. A. and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 604−613.
12. Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R., and Zhu, G. (1994) Measurement of homo- and heteronuclear J-couplings from quantitative J correlation, *Methods Enzymol. 239*, 79−105.
13. Wishart, D. S., and Nip, A. M. (1998) Protein chemical shift analysis: A practical guide, *Biochem. Cell Biol. 76*, 153−163.
14. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York, N. Y.
15. Duggan, B. M., Legge, G. B., Dyson, H. J., and Wright, P. E. (2001) SANE (Structure assisted NOE evaluation): An automated model-based approach for NOE assignment, *J. Biomol. NMR 19*, 321−329.
16. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
17. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (2002) *AMBER 7*, University of California, San Francisco, CA.
18. Tsui, V. and Case, D. A. (2000) Molecular simulations of nucleic acids using a generalized Born solvation model, *J. Am. Chem. Soc. 122*, 2489−2498.
19. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK−NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
20. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
21. Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators, *Nature 415*, 549−553.
22. Radhakrishnan, I., Pérez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions, *Cell 91*, 741−752.

BI048161T