Structural Diversity in p160/CREB-binding Protein Coactivator Complexes

Received for publication, January 10, 2006, and in revised form, March 14, 2006 Published, JBC Papers in Press, March 15, 2006, DOI 10.1074/jbc.M600237200

Lorna Waters†,1,2, Baigong Yue†,2, Vaclav Veverka†, Philip Renshaw†, Janice Bramham†, Sachiko Matsuda†, Thomas Frenkeli, Geoffrey Kelly†, Frederick Musket§, Mark Carr¶, and David M. Heery‡

From the †Department of Biochemistry, Henry Wellcome Building, University of Leicester, Lancaster Road, Leicester LE1 9HN, United Kingdom, ‡School of Pharmacy, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom, and §Medical Research Council Biomedical NMR Centre, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

Ligand-induced transcription by nuclear receptors involves the recruitment of p160 coactivators such as steroid receptor coactivator 1 (SRC1), in complex with histone acetyltransferases such as CREB-binding protein (CBP) and p300. Here we describe the solution structure of a complex formed by the SRC1 interaction domain (SID) of CBP and the activation domain (AD1) of SRC1, both of which contain four helical regions (C1, C2, C3, and C4) in CBP and So1, So2, So3, and So4 in SRC1. A tight four-helix bundle is formed between So1, C1, C2, and C3 that is capped by So4. In contrast to the structure of the AD1 domain of the related p160 protein ACTR in complex with CBP SID, the sequences forming So2 and So3 in SRC1 AD1 are not involved in the interface between the two domains but rather serve to position So4. Thus, although the CBP SID domain adopts a similar fold in complex with different p160 proteins, the topologies of the AD1 domains are strikingly different, a feature that is likely to contribute to functional specificity of these coactivator complexes.

The lysine acetyltransferase CBP interacts with a large number of nuclear proteins, many of which are transcription factors (1, 2). This is achieved through direct or indirect protein-protein interactions that are mediated by distinct structural domains within the CBP protein such as the KIX, CRD, CH1, CH3, bromodomain and SID domains. The protein-binding domains of CBP display partial specificity, having both distinct and overlapping binding partner profiles, which contributes to the phenomena of synergy and cross-talk between transcription factors (1).

The recruitment of CBP to target gene promoters/enhancers facilitates acetylation of histone N-terminal tails, leading to chromatin remodeling and enhanced gene expression. This has been demonstrated for nuclear receptors, which activate transcription of their target genes in response to ligand binding (3, 4). Ligand-bound receptors undergo a conformational change that stimulates their interaction with cofactors that contain functional LXXLL motifs, such as the p160 coactivators SRC1, TIF2, and ACTR (5, 6). Studies of steroid-regulated gene promoters have revealed that p160s and HAT proteins are among the first cofactors recruited in response to ligand (7, 8). Such temporally ordered recruitment of coactivators to promoters/enhancers is crucial for the sequential chromatin modification and remodeling events preceding transcription (4).

The efficacy of nuclear receptor/cofactor interaction is influenced by a number of determinants, including the precise sequence and number of LXXLL motifs, the sequences flanking core motifs, and other distal sequences (9–11). CBP contains three LXXLL motifs, although they mediate only weak direct interactions with estrogen, androgen, and progesterone receptors (9, 10). Thus, at least in the case of the steroid receptors, efficient recruitment of CBP/p300 and associated factors is achieved indirectly via the p160 proteins (12).

A number of studies have shown that recruitment of CBP/p300 proteins is facilitated by the p160 activation domain AD1, which acts as a potent transcriptional activator in mammalian and yeast cells (3, 12–17). The AD1 domain docks with the SRC1 interaction domain (SID) of CBP, located within the sequence 2058–2130 (12), and this interaction is essential for ligand-dependent transcription mediated by steroid receptors (12, 18). In addition to binding p160s, the SID also facilitates interactions of CBP with activation domains in other nuclear proteins, including the transcription factors Ets1, Ets2, p53, and IRF3, and viral activators such as E1A, KSHV IRF1, and Tax (19–22). Indeed, it has been shown that competition between such proteins for binding to the SID contributes to the negative cross-talk observed between different signaling pathways (22). Similarly, binding of viral proteins to the SID and other CBP/p300 domains not only facilitates viral gene transcription but may also down-regulate expression of host defense genes, through exclusion of host factors from CBP/p300 complexes.

The AD1/SID interaction is also important in MOZ-TIF2-associated leukemogenesis. MOZ-TIF2 is a fusion protein expressed in acute myeloid leukemia blasts containing the inv(8)(p11;q13) translocation (23). This fusion protein contains the C-terminal sequence of TIF2, including the AD1 domain, facilitating its interaction with CBP/p300, as demonstrated by in vitro interactions and in vivo by using fluorescence resonance energy transfer experiments (24). As a consequence of this interaction, CBP is mislocalized from promyelocytic leukemia bodies, and cellular levels of CBP are depleted, leading to a reduced transcriptional...
activity of CBP-dependent activators such as nuclear receptors and p53 (24). Consistent with this, AD1 integrity was found to be essential for transformation of hematopoietic progenitors by MOZ-TIF2 in vitro (24, 25) and for induction of acute myeloid leukemias by MOZ-TIF2 in mice (25). Therefore, understanding the structure of the CBP-p160 complexes has relevance to the etiology of acute myeloid leukemia.

Phenotypic differences in knock-out mice indicate that p160s have tissue-specific functions (reviewed in Ref. 26). Similarly, CBP and p300 proteins appear to have distinct roles in vivo, for example in hematopoiesis (27). Thus, the existence of different p160-CBP or p160-p300 complexes in vivo suggests they may have specific albeit partially redundant functions. For example, SRC1 and ACTR interact with thymine DNA glycosylase, whereas TIF2 does not. This is because of the presence of a tyrosine repeat motif (YXXY) in ACTR and SRC1, which is not conserved in the TIF2 sequence (28). Chromatin immunoprecipitation assays investigating cofactor recruitment at the promoter have indicated that the presence of one p160 can exclude recruitment of others (7), although the molecular basis of this is unknown. Another study provided evidence that CBP-p160 complexes are functionally distinct, as it was observed that different p160 combinations can be detected on androgen receptor target gene promoters, possibly through the formation of specific p160 heterodimer pairs (29). Thus, determination of the structures of different complexes will be essential to understand how such selectivity is achieved. In this study, we describe a high resolution solution structure of the complex formed between the SID domain of CBP and the AD1 domain of SRC1. By comparison to the CBP SID/ACTR AD1 structure (30), we show that although the structure of CBP SID is strikingly similar in both complexes, p160 proteins adopt distinct conformations that are likely to be important for their different biological functions.

**EXPERIMENTAL PROCEDURES**

**Protein Expression/Purification**—To express the CBP/SRC1 interaction domain complex in *Escherichia coli*, we obtained a modified pET22b dual expression vector, containing sequences encoding the ACTR AD1 and CBP SID domains, preceded by independent ribosome-binding sites (a gift from Peter Wright, Scripps Institute). To generate the CBP SID/SRC1 AD1 expression vector, the ACTR AD1 structure was removed by restriction digestion with NcoI and HindIII and replaced with an NcoI/HindIII-digested PCR fragment containing the coding region for residues 920–970 of SRC1 followed by a thrombin cleavage site (underlined) and a polyhistidine tag (KLVPRGSLEH). The mouse CBP SID-(2059–2117) and human SRC1 AD1-(920–970) dual expression vector was transformed into *E. coli* strain B834, which was used to produce either unlabeled, 15N-labeled, or 13C/15N-labeled samples of the CBP SID-SRC1 AD1 complex. For labeled samples, cells were grown in minimal media containing 0.6 g/liter [15N]ammonium sulfate and 2 g/liter [13C]glucose as the sole nitrogen and/or carbon sources. The cells were harvested 3 h post-induction, resuspended in 20 mM sodium phosphate, 100 mM sodium chloride, and 5 mM imidazole buffer (pH 7.0). Lysis was achieved by sonication. Insoluble material was removed by high speed centrifugation, and the cleared lysate was filtered through a 0.2-µm filter prior to chromatography. The histidine-tagged CBP SID-SRC1 AD1 complex was affinity-purified on a HiTrap chelating HP column (Amersham Biosciences) charged with Ni2+ and eluted with a linear gradient of imidazole (5–250 mM). Fractions containing the complex were pooled and subjected to a final purification by gel filtration on a Superdex 75 16/60 pre-packed column (Amersham Biosciences). Typical yields of the >95% pure complex were 10 mg/liter.

**Reverse Phase HPLC**—Reverse phase HPLC analysis was performed on a C4 column. The individual polypeptides were resolved by applying a two-step linear gradient of acetonitrile (1.6–80%), and the relative amount of each was quantified by absorbance at 215 nm. Eluted peaks were identified by mass spectrometry.

**Circular Dichroism Spectroscopy**—Far UV CD spectra were obtained from 0.4 mM samples of the CBP SID-SRC1 AD1 complex in a 20 mM sodium phosphate, 100 mM sodium chloride buffer (pH 7.0). Samples were placed in a 0.1-mm path length cell, and spectra were recorded from 180 to 250 nm at a resolution of 1 nm and a scan speed of 20 nm/min, with each spectrum representing the average of 10 accumulations. Prior to analysis, CD spectra were corrected for buffer absorbance, and the raw data were converted to molar CD per residue.

**NMR Spectroscopy**—NMR spectra were acquired from 0.35-mL samples of 1.5 mM CBP SID-SRC1 AD1 complex in a 20 mM sodium phosphate, 100 mM sodium chloride, 10 µM EDTA, and 0.02% (w/v) sodium azide buffer (pH 7.0), containing either 10% D2O, 90% H2O, or 100% D2O as appropriate. All NMR data were acquired at 25 °C on either an 800-MHz Varian Inova or a 600-MHz Bruker Avance spectrometer. The two-dimensional and three-dimensional spectra recorded to obtain sequence-specific assignments for CBP SID and SRC1 AD1 in complex were as follows: 1H TOCSY with mixing times of 40 and 55 ms (31) and NOESY with an NOE mixing time of 100 ms (32); 15N/1H HSQC; TOCSY-HSQC with a mixing time of 50 ms and NOESY-HSQC with an NOE mixing time of 100 ms (33); 13C/1H HCCH-TOCSY with a mixing time of 20 ms (34), HMBC-NOESY with an NOE mixing time of 100 ms (35); and 15N/13C/1H HNCA/CB (36) and CBCA(CO)NH (37). Typical acquisition times in F1 and F2 for the three-dimensional experiments were 11–13 ms for 15N, 7.5–9.5 ms for 13C, and 15 ms for 1H, and an acquisition time of 75 ms in F3 (1H). The majority of the three-dimensional spectra were collected over ~88 h, two-dimensional 1H experiments over 8.5–24 h, and 1N/1H HSQC spectra over about 30 min. Typical acquisition times in two-dimensional experiments were either 70 (1H) or 35 ms (1H) in F1 and 250 ms in F2 (1H). The WATERGATE method (38) was used to suppress the water signal when required. The three-dimensional NMR data were processed using NMRPipe (39) with linear prediction used to extend the effective acquisition times by up to 1.5–2-fold in F1 and F2. The spectra were analyzed using the XEASY package (40).

**Structure Calculations**—A family of converged CBP SID/SRC1 AD1 structures was determined in a two-stage process using the program CYANA (41). Initially, the combined automated NOE assignment and structure determination protocol (CANDID) was used to automatically assign both the intra- and intermolecular NOE cross-peaks identified in three-dimensional 15N- and 13C-edited NOESY spectra. This approach provides a completely unbiased assignment of the NOE peaks, in which all peaks are evaluated as either intra- or intermolecular in origin. Subsequently, several cycles of simulated annealing combined with redundant dihedral angle constraints were used to produce the final converged CBP SID/SRC1 AD1 structures (42). The input for the CANDID stage consisted of essentially complete 15N, 13C, and 1H resonance assignments for the nonexchangeable groups in the CBP SID-SRC1 AD1 complex, two manually picked three-dimensional NOE peak lists corresponding to all NOEs involving amide protons (1179) and all NOEs between aliphatic protons (2371), and one manually picked two-dimensional NOE peak list corresponding to all NOEs involving aromatic side chain protons (127). In addition, the CANDID stage included
80 torsion angle constraints for CBP SID and 64 torsion angle constraints for SRC1 AD1 obtained from analysis of chemical shift data using TALOS (43). Hydrogen bond constraints were also added for regions predicted to contain regular α-helical secondary structure by both the NOE and chemical shift data, which corresponds to residues Ala2087–Leu2097, Gin2098–Lys2107, and Ala2109–Ala2117 of CBP SID and Glu929–Ser941 of SRC1 AD1. The peak lists were prepared using XEASY, and the intensities of peaks were obtained using the "interactive integration" routine (40). CANDID calculations were carried out using the default parameter settings in CYANA 1.0.6, with the upper limit for NOE-derived distance constraints set at 5.5 Å and chemical shift tolerances set to 0.02 ppm (direct and indirect $^1H$) and 0.3 ppm ($^{15}N$ and $^{13}C$). The final converged CBP SID/SRC1 AD1 structures were produced from 100 random starting coordinates using a standard torsion angle-based simulated annealing protocol combined with 5 cycles of redundant dihedral angle constraints (44, 45). The calculations were based upon 1855 nonredundant NOE-derived upper distance limits (maximum value 6.0 Å), assigned to unique pairs of protons using CANDID, 142 $\Phi$ and $\Psi$ torsion angle constraints derived from TALOS, and 104 hydrogen bond constraints for helical residues. Analysis of the family of structures obtained was carried out using the programs CYANA and MOLMOL (41, 46).

Sequence-specific Assignments—Sequence-specific backbone resonance assignments (N, NH, $^1C\beta$, and $^1C\gamma$) were obtained for the CBP SID-SRC1 AD1 complex from the identification of intra- and inter-residue connectivities in HNCACB, CBCA(CO)NH, and $^{15}N/H$ NOESY-HSQC spectra. Assignments were then extended to the side chain signals using correlations observed primarily in $^{15}N/H$ TOCSY-HSQC and $^{13}C/H$ HCCCH-TOCSY, with additional supporting evidence provided by $^{15}N/H$ NOESY-HSQC and $^{13}C/H$ HMBC-NOESY spectra where required.

RESULTS AND DISCUSSION

Expression and Purification of the SRC1-CBP Complex—We previously used yeast two-hybrid and in vitro pulldown experiments to define the minimal sequences required for interaction of the CBP SID and SRC1 AD1 domains (12). A modified pET22b dual expression vector (30), containing sequences encoding the SRC1 AD1-(920–970) and CBP SID-(2059–2117) domains, was used to coexpress these polypeptides in E. coli. The addition of a polyhistidine tag at the C terminus of the AD1 domain facilitated purification of the complex, as described under “Experimental Procedures.” A highly purified complex containing polypeptides of the expected size was obtained, as determined by SDS-PAGE (Fig. 1A). Far UV CD spectra acquired for the CBP SID-SRC1 AD1 complex indicated that the complex was predominantly helical, having characteristic negative ellipticity peaks at -209 and 221 nm (Fig. 1B). Analysis of the spectra using the CDPro software package (47) indicated that the complex contained approximately equal amounts of CBP SID and SRC1 AD1 domains.

FIGURE 1. Purification and characterization of the CBP SID-SRC1 AD1 complex. A, 15% SDS-polyacrylamide gel showing separation of whole cell lysate from E. coli B834 expressing CBP SID-SRC1 AD1 complex (lane 2). The His$_6$-tagged complex was purified by affinity chromatography. Lanes 3–12 show fractions containing the purified complex after gel filtration. Molecular weight markers are shown in lane 1. B, far UV circular dichroism spectrum acquired for the CBP SID-SRC1 AD1 complex. C, reverse phase HPLC of the purified CBP SID-SRC1 AD1 complex. The areas under the two peaks were used to calculate the ratio of the CBP SID domain to SRC1 AD1 domain in the purified complex. D, $^{15}N/H$ HSQC spectrum of the CBP SID-SRC1 AD1 complex. The assignments of the signals from backbone amide groups in both domains are indicated by residue type and number, with the overlapped region between 7.85 and 8.35 ppm in $^1H$ and 118.5 and 121.5 ppm in $^{15}N$ shown in the expanded region to the right of the complete spectrum. For clarity, residues are numbered according to their position in the domain. Thus, CBP residues Pro2059–Pro2117 are numbered P2–P19, and SRC1 residues Pro920–Ser941 are numbered P930–P953, with the C-terminal linker to the His tag numbered 354–361. Assignments obtained for a number of side chain NH$_2$ groups are indicated on the spectrum.
Structure of the SRC1-CBP Complex

The solution structure of the SRC1-CBP complex was determined to high precision, as evident from the superposition of the protein backbone of the family of converged structures (best fit for residues 2063–2113 of CBP SID and 928–963 of SRC1 AD1; see Fig. 2A) and reflected in low root mean square deviation (r.m.s.d.) values to the mean structure for both the backbone (0.49 ± 0.10 Å) and all heavy atoms (0.95 ± 0.09 Å). The backbone topology of the complex is composed primarily of eight helices linked by turns and loops, including four α-helices in the CBP SID domain (Ca1, Ser2066–Lys2076; Ca2, Glu928–Ser939; Ca3, Pro909–Thr920; and Ca′3, Tyr919–Asn923) and three α-helices in the SRC1 AD1 domain (Ser1, Glu929–Ser941; So2, Glu945–Leu949; and So3, Ile957–Gln962). SRC1 AD1 also contains a short 3_10-helix (So2, Glu959–Leu955). The total helical content of the complex is just over 50%, in close agreement with the CD analysis (54% helical structure; see Fig. 1B).

The CBP SID and SRC1 AD1 domains are intimately associated in the complex, with Ca1, Ca2, Ca3, and So1 forming a four-helix bundle (Fig. 2B). The overall interface between the two domains is substantial, corresponding to a solvent-inaccessible surface of 1019 Å² on CBP SID and 1088 Å² on SRC1 AD1. Helices So2′ and So2 pack together on one side of the complex to stabilize the corner region of the “L”-shaped domain and are not involved in the interface with CBP SID. In contrast, So3 serves to cap the four-helix bundle region and forms contacts with Ca2, Ca3, Ca′3, and the two serine residues in the PSSP turn of CBP SID. The first two helices of CBP SID (Ca1 and Ca2) lie almost antiparallel to each other and are separated by a well defined five-residue turn consisting of the sequence SPSSP. The N-terminal portion of Ca2 contains a stretch of five glutamines preceding the more hydrophobic C-terminal half of the helix. Ca3 is inclined away from Ca1 and Ca2 exposing a hydrophobic groove between Ca1 and Ca3, which accommodates So1 of the SRC1 AD1 domain. Helices Ca3 and Ca′3 are separated by a short region of irregular structure and resemble a single bent helix. The stabilizing interactions between Ca1 and Ca2 and between Ca2 and Ca3 appear to involve no ionic or hydrogen bonds but rely on favorable van der Waals contacts, primarily involving residues Ala970, Leu971, and Thr974 from Ca1; Gln973, Val976, Ile998, and Leu999 from Ca2; and Leu997, Phe1001, and Thr1006 from Ca3. Asn997 also makes van der Waals interactions with both Ca2 and Ca3.

The SRC1 AD1 domain also appears to be primarily stabilized by van der Waals interactions, involving residues Ala981 and Gln985 from So1.

Table 1. The family of converged CBP SID-SRC1 AD1 complex structures, together with the NMR constraints, have been deposited in the Protein Data Bank (code 2C52).

| TABLE 1 |
| NMR constraints and structural statistics for the CBP SID-SRC1 AD1 complex |

(a) No. of constraints used in the final structural calculations

| Type of NOE | No. of constraints |
|-------------|--------------------|
| Intrareidue NOEs | 328 |
| Sequential NOEs (i,i+1) | 514 |
| Medium range NOEs (i,i+4) | 643 |
| Intramolecular range NOEs (i,i+5) | 166 |
| Torsion angles | 108 |
| Hydrogen bonding | 104 |

(b) Maximum and total constraint violations in 37 converged CBP SID/SRC1 AD1 structures

| Constraint type | Maximum violation (Å) | Total violation (Å) |
|----------------|------------------------|---------------------|
| Upper distance limits | 0.40 ± 0.05 | 20.7 ± 1.84 |
| Lower distance limits | 0.24 ± 0.08 | 1.5 ± 0.24 |
| van der Waals contacts | 0.28 ± 0.04 | 14.8 ± 1.55 |
| Torsion angle violations (°) | 3.44 ± 0.64 | 33.8 ± 6.53 |
| Average CYANA target function (R²) | 6.17 ± 0.95 |

(c) Structural statistics for the family of converged CBP SID/SRC1 AD1 structures

- Residues within allowed region of Ramachandran plot (%): 98 (0.24 Å, 14.8°, 71°)
- No. of constraints used in the final structural calculations: 54% (±7.3%) helical, 20% (±6.1%) turn, and 25% (±5.7%) random coil secondary structure.

The CANDID protocol was effective in determining unique assignments for the NOEs identified in the three-dimensional 15N- and 13C-edited NOESY and the aromatic to aliphatic region of the two-dimensional HSQC spectrum (Fig. 1C).

Sequence-specific Assignments and Structure Calculation for the CBP-SRC1 Complex—Very comprehensive sequence-specific resonance assignments were obtained for the CBP SID-SRC1 AD1 complex despite the relatively poor dispersion observed in spectra, which is illustrated by the HSQC spectrum (Fig. 1D). For example, backbone amide assignments were obtained for all non-proline residues in the complex except as follows: Asn2060, Arg2061, and Gln2117 in CBP SID; Asn927 in SRC1 AD1; and Pro2059 and Thr2106, and Cys2107 in CBP SID-SRC1 AD1 complex (Fig. 1E).

The CBP SID-SRC1 AD1 complex structures were determined using NMR constraints and structural statistics for the family of converged CBP SID-SRC1 AD1 structures (see Table 1). The converged structures (best fit for residues 2063–2113 of CBP SID and 928–963 of SRC1 AD1; see Fig. 2A) and reflected in low root mean square deviation (r.m.s.d.) values to the mean structure for both the backbone (0.49 ± 0.10 Å) and all heavy atoms (0.95 ± 0.09 Å). The backbone topology of the complex is composed primarily of eight helices linked by turns and loops, including four α-helices in the CBP SID domain (Ca1, Ser2066–Lys2076; Ca2, Glu928–Ser939; Ca3, Pro909–Thr920; and Ca′3, Tyr919–Asn923) and three α-helices in the SRC1 AD1 domain (Ser1, Glu929–Ser941; So2, Glu945–Leu949; and So3, Ile957–Gln962). SRC1 AD1 also contains a short 3_10-helix (So2, Glu959–Leu955). The total helical content of the complex is just over 50%, in close agreement with the CD analysis (54% helical structure; see Fig. 1B).

The CBP SID and SRC1 AD1 domains are intimately associated in the complex, with Ca1, Ca2, Ca3, and So1 forming a four-helix bundle (Fig. 2B). The overall interface between the two domains is substantial, corresponding to a solvent-inaccessible surface of 1019 Å² on CBP SID and 1088 Å² on SRC1 AD1. Helices So2′ and So2 pack together on one side of the complex to stabilize the corner region of the “L”-shaped domain and are not involved in the interface with CBP SID. In contrast, So3 serves to cap the four-helix bundle region and forms contacts with Ca2, Ca3, Ca′3, and the two serine residues in the PSSP turn of CBP SID. The first two helices of CBP SID (Ca1 and Ca2) lie almost antiparallel to each other and are separated by a well defined five-residue turn consisting of the sequence SPSSP. The N-terminal portion of Ca2 contains a stretch of five glutamines preceding the more hydrophobic C-terminal half of the helix. Ca3 is inclined away from Ca1 and Ca2 exposing a hydrophobic groove between Ca1 and Ca3, which accommodates So1 of the SRC1 AD1 domain. Helices Ca3 and Ca′3 are separated by a short region of irregular structure and resemble a single bent helix. The stabilizing interactions between Ca1 and Ca2 and between Ca2 and Ca3 appear to involve no ionic or hydrogen bonds but rely on favorable van der Waals contacts, primarily involving residues Ala970, Leu971, and Thr974 from Ca1; Gln973, Val976, Ile998, and Leu999 from Ca2; and Leu997, Phe1001, and Thr1006 from Ca3. Asn997 also makes van der Waals interactions with both Ca2 and Ca3.

The SRC1 AD1 domain also appears to be primarily stabilized by van der Waals interactions, involving residues Ala981 and Gln985 from So1.
Lys$^{943}$ from the unstructured region between Sa1 and Sa2$, Thr^{946}$ from Sa2$, Ala^{949}$ that links Sa2$ to Sa2$, and Gly$^{956}$ that links Sa2 to Sa3$. However, two potential salt bridges are formed between residues Glu$^{945}$ and Lys$^{950}$ and between Asp$^{952}$ and Lys$^{959}$: The binding of Sa1 in the hydrophobic groove situated between Ca1 and Ca3 is stabilized by van der Waals interactions. The favorable contacts between the N-terminal helices of CBP SID and SRC1 AD1 mainly involve interactions between buried side chains found in the leucine-rich motifs of Ca1 (LXXLXXL corresponding to Leu$^{906}$, Leu$^{1071}$, Leu$^{1107}$, Leu$^{1107}$) and Sa1 (LXXLXXFL corresponding to Leu$^{932}$ and Leu$^{936}$). These motifs resemble the LXXL and /LXXL/H/LXXXL motifs that mediate the interaction of coactivators and corepressors with nuclear receptors. Additional favorable van der Waals interactions are found between residues Gln$^{935}$ and Phe$^{939}$ from Sa1 and the leucine residues of Ca1. Similarly, stabilizing contacts occur between Sa1 and Ca3$ involving the leucine-rich motif (Leu$^{935}$, Leu$^{936}$, and Leu$^{940}$) and residues Glu$^{929}$ and Val$^{937}$ of Sa1, and residues Gln$^{938}$, Leu$^{939}$, Ala$^{940}$, Phe$^{944}$, and Gln$^{949}$ of Ca3 in CBP. A potential salt bridge between Lys$^{970}$ and Asp$^{944}$ may also stabilize this region of the complex.

The fourth helix of SRC1 AD1 (Sa3) fits into a second hydrophobic groove in CBP SID located between the well defined SSPP turn linking Ca1 and Ca2 and the C terminus of Ca3. This interaction is stabilized by hydrophobic contacts involving Leu$^{955}$, Ile$^{957}$, Leu$^{960}$, Val$^{961}$, Gln$^{962}$, and Gly$^{964}$ of SRC1 AD1 and Ser$^{2078}$, Ser$^{1080}$, Gln$^{1084}$, Arg$^{2105}$, Lys$^{2108}$, and Tyr$^{2109}$ of CBP SID. The five C-terminal residues of SRC1 AD1 appear to loop back toward Sa3 and Ca3$, leading to some van der Waals contacts between Leu$^{960}$ of SRC1 AD1 and Tyr$^{2109}$ of CBP SID. However, this loop is primarily stabilized by intramolecular van der Waals interactions involving Gln$^{962}$ and residues within a C-terminal linker sequence originating from the expression vector. In the absence of the linker sequence, contacts between Leu$^{960}$ and Tyr$^{2109}$ may not occur, and the C-terminal region of SRC1 AD1, including Sa3$, may make additional contacts within the complex, or with other domains/proteins.

Some features of the NMR data obtained for the CBP SID-SRC1 AD1 complex clearly show the presence of a second minor conformational state, which is in exchange with the structure reported here (Fig. 1D). This is clearly indicated by the presence of cross-peaks between signals from backbone amide groups in $^{15}$N/$^{1}$H TOCSY-HSQC spectra, which arise through chemical exchange processes (48). This also results in exchange broadening of a significant number of backbone amide signals. Interestingly, only a few of the TOCSY exchange peaks involve residues in CBP SID, whereas almost all of the residues from Leu$^{936}$–Val$^{941}$ in SRC1 AD1 are affected, which corresponds to the six C-terminal residues of Sa1 through to the end of Sa3$. This suggests that the SRC1 AD1 domain has a degree of conformational instability, even when bound to CBP SID, which may facilitate the rapid dissociation of the complex. However, we do not exclude the possibility that other protein/protein interactions may stabilize complexes of the full-length proteins.

Comparison of CBP-SRC1 and CBP-ACTR Complexes—The solution structure of CBP SID in complex with the AD1 domain from ACTR has been reported previously (30). Comparison of the two complexes revealed that they have both conserved and distinct structural features (Fig. 3, A and B). The SID adopts similar secondary and tertiary structures, when bound to different AD1 domains (Fig. 3, A and B), as highlighted by the superposition of the SID polypeptide backbone atoms (residues Ile$^{2063}$, Pro$^{2065}$–Ala$^{2067}$, Asp$^{2070}$–Ser$^{2079}$, and Gln$^{2085}$–Lys$^{2108}$), which yields an r.m.s.d. value of 1.79 Å (Fig. 3C). The most notable difference between the CBP SID polypeptides in the two complexes is the conformation of the sequence composed of five glutamine residues (2082–2086). In the CBP-SRC1 complex this region forms the N terminus of an extended Ca2 helix (Fig. 3, A and C), whereas in the CBP-ACTR complex the three proximal glutamine residues were pro-

FIGURE 2. Solution structure of the CBP SID-SRC1 AD1 complex. A, best fit superimposition of the family of 37 converged NMR structures obtained for the CBP SID-SRC1 AD1 complex. The CBP SID domain is shown in blue and SRC1 AD1 in red. B, schematic (ribbon) representation of the backbone topology of the two polypeptides in the complex, using the same orientation as in A, and demonstrating the tight 4-helix bundle conformation. The four helices of CBP SID (Ca1, Ca2, Ca3, and Ca3$)$ and SRC1 AD1 (Sa1, Sa2$, Sa2$, and Sa3) are indicated, along with the N and C termini of both domains.
posed to form part of a solvent-exposed loop (designated poly(Q); Fig. 3, B and C). This difference may be due to the paucity of NMR signal assignments for residues in this region in the CBP-ACTR complex (30), resulting in fewer NMR constraints on the conformation rather than a true structural difference. Another difference between the CBP SID domains is Cα3, which is 17 residues long (Pro2095–Ala2111) in the CBP/ACTR structure, although in the CBP-SRC1 complex this region forms two shorter helices (Cα3, Pro2095–Thr2106, Caα3′, Tyr2109–Asn2113) resembling a kinked helix. The Caα3′ helix is inclined away from the rest of the four-helix bundle, which is not observed in the ACTR-AD1 complex.

Structural similarity between p160 AD1 domains in the two complexes is confined to the N-terminal helices Aα1 and So1, which form four-helix bundles with the CBP SID (Fig. 3, A and B). The remaining sequences in the AD1 domains adopt strikingly distinct topologies as highlighted by superimposing the AD1 polypeptide structures (Fig. 3D). This is also reflected in the significant chemical shift differences for NMR signals from equivalent residues in the two AD1 domains and also

FIGURE 3. Comparison of CBP SID/SRC1 AD1 and CBP SID-ACTR AD1 complexes. A and B, equivalent views of the backbone topology of the CBP SID-SRC1 AD1 and CBP SID-ACTR AD1 complexes, respectively, which were obtained by superimposing the CBP SID domain from both complexes (residues Ile2063, Pro2065–Ala2067, Asp2070–Ser2079, and Gln2085–Lys2108). The CBP SID domain is shown in blue; SRC1 AD1 is shown in red and ACTR AD1 in green. C, comparison of the backbone topologies of the two CBP SID domains, which were overlaid on the same residues as in A and B. The CBP SID domains from the SRC1 AD1 and ACTR AD1 complexes are shown in blue and cyan, respectively. D, comparison of the backbone folds of the AD1 domains of SRC1 (red) and ACTR (yellow), which were overlaid on residues Ile2063, Pro2065–Ala2067, Asp2070–Ser2079, Gln2085–Lys2108 of CBP-SID and residues Asp928–Ser941 and Asp1044–Ser1057 of SRC1 and ACTR, respectively.
for CBP SID residues that make distinct contacts with AD1 in the two complexes, in particular, residues in the Ca3/Ca3’ region (data not shown). Although SRC1 AD1 contains three further helical regions (residues Sa2’, Glu945–Leu948, Sa2, Glu950–Ser954, and Sa3, Ile957–Gln962), ACTR AD1 contains only two additional helices (Asa2, Leu1064–Leu1071, and Asa3, Ile1073–Gln1079) (Fig. 3D). These distinct topologies are somewhat unexpected given that SRC1 and ACTR share 53% sequence identity and 67% sequence similarity within the AD1 domain. In the CBP SID-ACTR AD1 complex, Asa2 and Asa3 wrap around Ca3 making extensive contacts with this helix, whereas Asa2 occupies a hydrophobic groove between the proposed poly(Q) loop and Ca3 (Fig. 3B). In the CBP SID-ACTR AD1 complex, only the C-terminal residue (Leu1071) of Sa2 (corresponding to Leu1072 in Asa2) makes any contact with the SID domain (Fig. 3A), instead, Sa3 fills the groove between the PSSP turn and Ca3, effectively capping the four-helix bundle (Fig. 3A). This contrasts with the positions of corresponding helix of ACTR (Asa3), which packs against the adjacent hydrophobic face of Ca3 (Fig. 3B).

The CBP SID-ACTR AD1 complex was reported to contain a salt bridge between residues Asp944 in Asa2 and Arg2105 in Ca3, coordinating a hydrogen-bonding network involving Asp1068 and Arg2105 in Ca3, coordinating a hydrogen-bonding network involving Asp944 and Arg2105 and Asp1068 and Tyr2109. This buried charged interaction has been proposed to be important for the specificity of the CBP SID/ACTR AD1 interaction, although it does not contribute directly to the stability of the complex (49). Both of these aspartate residues are conserved in SRC1 AD1 (Asp944 and Asp952); however, they are spatially distant from the Ca3 helix. Asp944 is present in the loop between Sa1 and Sa2’, and Asp952 is contained within the Sa2 helix (Fig. 4C). However, it is possible that Asp944 forms a salt bridge with Lys2076 from Ca1. Similarly, an aspartate residue (Asp965) near the C terminus of SRC1 AD1 may form a salt bridge with Arg2105, which is involved in van der Waals contacts between Ca3 and Sa3. As would be expected from the tertiary structures of their respective complexes, the solvent-inaccessible surface areas of CBP SID (1681 Å²) and ACTR AD1 (1839 Å²) domains are greater than those observed for the complex of CBP SID (1019 Å²) and SRC1 AD1 (1088 Å²).

In summary, the backbone structure of the CBP SID and the first helix of AD1 are very similar in both complexes, as reflected in a backbone atom r.m.s.d. value of 2.05 Å (calculated by superimposition of residues Leu2063, Pro2065-Ala2067, Arg2070-Ser2079, and Gln2085-Lys2108 of CBP SID and residues Asp928–Ser941 and Asp1044–Ser1057 of SRC1 AD1 and ACTR-AD1, respectively). The remainder of the p160 AD1 domains adopt very distinct folds, although the C-terminal helices of AD1 appear to be essential to stabilize the complex.

Structural Flexibility of the CBP SID Permits Complex Formation with Multiple Partners—The NMR structure of the CBP SID in isolation (also termed IBiD because of its interaction with IRF3) revealed it to have significant α-helical content (19), although the conformation of the isolated CBP SID has been proposed to be consistent with that of a molten
globule (50). Isolated AD1 polypeptides have little if any intrinsic structure (50) but appear to undergo induced folding in complex with CBP SID (30). An important question is whether short conserved structural motifs promote similar modes of binding of different proteins to CBP, or whether different complexes adopt drastically different conformations that influence the overall structure of CBP/p300 proteins, which in turn influences combinatorial complex formation.

This study reveals structural similarity within the α1 helix of two p160 AD1 domains, which contributes to a four-helix bundle in both the CBP-SRC1 and CBP-ACTR complexes. We have shown previously that the LXXLXXXXL sequence motif within α1 is conserved in other SID-binding proteins (see Fig. 4A), and that disruption of this sequence in Ets2 and E1A abrogates their binding to the CBP SID (22). Fusion of the SRC1 AD1 sequence to a GAL4 DNA binding domain permits very potent activation of a GAL reporter gene in mammalian cells, through recruitment of endogenous CBP and p300 (12). As shown in Fig. 4B, replacement of the Leu932, Leu933, and Leu936 residues in Saα1 with alanines resulted in loss of AD1 transcriptional activity. This highlights the requirement of the Saα1 for interaction with full-length CBP/p300 in vivo.

The C-terminal helix (α3) of the SRC1 and ACTR AD1 domains also plays a critical role in formation of the two complexes with CBP SID as discussed above. Interestingly, the sequence of the Saα3 helix of SRC1 AD1 (57/5DKLVL96/1) is divergent from that in ACTR and TIF2 (IPELV) (Fig. 4A). The IPCLV sequence in ACTR and TIF2 resembles the LPXL motif that has been shown to mediate the interaction of CITED and HIF1α transcription factors with the CH1 domain of CBP (51). Mutagenesis of the SRC1 AD1 domain in the context of GAL4-AD1 revealed that the D958A or L960A mutations resulted in almost complete loss of reporter activation, whereas K959A retained almost full transactivation function (Fig. 4B). In addition, reporter activation by the V961A mutant was reduced by 70%, suggesting it is important but not essential for CBP/p300 recruitment (Fig. 4B). These results confirm the importance of Saα3 and key residues within that helix for CBP/p300 recruitment.

Replacement of the conserved aspartate residues (D944A and D952A) also resulted in loss of reporter activation by GAL4-AD1 (Fig. 4B). Thus, as suggested for the CBP-ACTR complex (30, 49), salt bridge formation may be important in stabilizing the higher order folding of CBP-p160 complexes. There is evidence to suggest that SRC1 sequences outside of the AD1 domain may influence the interaction with CBP. The signaling molecule 8-bromo-cAMP induces mitogen-activated protein kinase-dependent phosphorylation of SRC1 at Thr1179 and Ser1185, which leads to increased ligand-independent transcriptional activity of the proges-

trical changes in SRC1 that effect the affinity of the AD1 region for CBP.

We have indicated previously that several CBP SID-binding proteins contain low level sequence homology in the regions of SRC1 AD1 corresponding to Saα1, and to a lesser extent Saα3, that may indicate conserved modes of binding to CBP. An alignment of the sequences of interaction domains of several CBP SID-binding proteins is shown in Fig. 4A. The transactivation domain of human T cell leukemia Tax protein contains a short sequence (51/2YTNIPISL51/9) that is required for binding to CBP SID (21). This sequence may form an amphipathic α-he-

lix analogous to Saα3 of SRC1 AD1 and play an equivalent role in stabilizing CBP/Tax interactions (Fig. 4A).

REFERENCES

1. Goodman, R. H., and Smolik, S. (2000) Genes Dev. 14, 1553–1577
2. Giordano, A., and Avantaggiati, M. L. (1999) J. Cell. Physiol. 181, 218–230
3. Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Glass, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1996) Cell 85, 403–414
4. Perissi, V., and Rosenfeld, M. G. (2004) Nat. Rev. Mol. Cell Biol. 6, 542–554
5. Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997) Nature 387, 733–736
6. Torchia, J., Rose, D. W., Inostroza, J., Kamei, Y., Westin, S., Glass, C. K., and Rosenfeld, M. G. (1997) Nature 387, 677–684
7. Metivier, R., Penot, G., Hubner, M. R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003) Cell 115, 751–763
8. Shang, Y., Hu, X., DiRenzo, J., Lazar, M. A., and Brown, M. (2000) Cell 103, 843–852
9. Heery, D. M., Hoare, S., Hussain, S., Parker, M. G., and Shepperd, H. (2001) J. Biol. Chem. 276, 6695–6702
10. Coulthard, V. H., Matsuda, S., and Heery, D. M. (2003) J. Biol. Chem. 278, 19942–19951
11. Plevin, M. J., Mills, M. M., and Ikura, M. (2005) Trends Biochem. Sci. 30, 66–69
12. Sheppard, H. M., Harries, J. C., Hussain, S., Bevan, C., and Heery, D. M. (2001) Mol. Cell. Biol. 21, 39–50
13. Chen, H., Lin, R. J., Schiltz, R. L., Chakravarti, D., Nagy, L., Privalsky, M. L., Nakatani, Y., and Evans, R. M. (1997) Cell 90, 569–580
14. Voegel, J., Heine, M. J., Tini, M., Vivat, V., Chambon, P., and Gronemeyer, H. (1998) EMBO J. 17, 507–519
15. Kalkhoven, E., Valentine, J. E., Heery, D. M., and Parker, M. G. (1998) EMBO J. 17, 232–243
16. McNerney, E. M., Rose, D. W., Flynn, S. E., Westin, S., Mullen, T. M., Krones, A., Inostroza, J., Torchia, J., Nolte, R. T., Assa-Munt, N., Milburn, M. V., Glass, C. K., and Rosenfeld, M. G. (1998) Genes Dev. 2, 3357–3366
17. Sheppard, H. M., Matsuda, S., Harries, J. C., Kindle, K. B., and Heery, D. M. (2001) J. Mol. Endocrinol. 30, 411–427
18. Kim, M. Y., Hsiao, S. J., and Kraus, W. L. (2001) EMBO J. 20, 6084–6094
19. Lin, C. H., Hare, B. J., Wagner, G., Harrison, S. C., Mantiatis, T., and Fraenkel, E. (2001) Mol. Cell 8, 581–590
20. Livingood, J. A., Scoggin, K. E. S., Van Orden, K., McNelly, S. J., Edayathumangalam, R., Layhoun, P. J., and Nyborg, J. K. (2002) J. Biol. Chem. 277, 9054–9061
21. Scoggin, K. E. S., Uloa, A., and Nyborg, J. K. (2001) Mol. Cell. Biol. 21, 5250–5253
22. Matsuda, S., Harries, J. C., Viskaduraki, M., Troke, P. J. F., Kindle, K. B., Ryan, C. M., and Heery, D. M. (2004) J. Biol. Chem. 279, 14055–14064
23. Carapezza, M., Aguilar, R. C., Goldman, J. M., and Cross, N. C. (1998) Blood 1, 3127–3133

Acknowledgments—We are grateful to Peter Wright for the gift of the dual expression plasmid pET22B. We thank Jonas Emsley and Paul McEwan for useful discussions and Sharad Mistry for technical assistance.
24. Kindle, K. B., Troke, P. J., Collins, H. M., Matsuda, S., Bossi, D., Belloci, C., Kalkhoven, E., Salomoni, P., Pelicci, P. G., Minucci, S., and Heery, D. M. (2005) Mol. Cell. Biol. 25, 988–1002
25. Deguchi, K., Ayton, P. M., Carapeti, M., Kutow, J. L., Snyder, C. S., Williams, I. R., Cross, N. C., Glass, C. K., Cleary, M. L., and Gilliland, D. G. (2003) Cancer Cell 3, 259–271
26. Xu, J., and Li, Q. (2003) Mol. Endocrinol. 17, 1681–1692
27. Kung, A. L., Rebel, V. I., Bronson, R. T., Ch’ng, L.-E., Sieff, C. A., Livingston, D. M., and Yao, T.-P. (2000) Genes Dev. 14, 272–277
28. Lucey, M. J., Chen, D., Phoenix, F., Lopez-Garcia, J., Hart, S. M., Al-Jehani, R., Alao, J. P., White, R., Kindle, K. B., Losson, R., Chambon, P., Parker, M. G., Schar, P., Heery, D. M., Buluwela, L., and Ali, S. (2005) Nucleic Acids Res. 33, 6393–6404
29. Zhang, H., Yi, X., Sun, X., Yin, N., Shi, B., Wu, H., Wang, D., Wu, G., and Shang, Y. (2004) Genes Dev. 18, 1753–1765
30. Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Nature 415, 549–553
31. Braunschweiler, L., and Ernst, R. R. (1983) J. Magn. Reson. 53, 521–528
32. Macura, S., and Ernst, R. R. (1980) Mol. Physiol. 41, 95–117
33. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) J. Am. Chem. Soc. 111, 1515–1517
34. Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) J. Magn. Reson. 88, 425–431
35. Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., and Fesik, S. W. (1990) J. Magn. Reson. 86, 210–216
36. Wittekind, M., and Mueller, L. (1993) J. Magn. Reson. 101, 201–205
37. Grzesiek, S., and Bax, A. (1993) J. Biomol. NMR 3, 185–204
38. Piotto, M., Saudek, V., and Sklenar, V. (1992) J. Biomol. NMR 2, 661–665
39. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
40. Bartels, C., Xia, T.-H., Billetter, M., Güntert, P., and Wüthrich, K. (1995) J. Biomol. NMR 5, 1–10
41. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) J. Mol. Biol. 319, 209–227
42. Güntert, P., and Wüthrich, K. (1991) J. Biomol. NMR 1, 447–456
43. Cornilescu, G., Delaglio, F., and Bax, A. (1999) J. Biomol. NMR 13, 289–302
44. Carr, M. D., Bloemink, M. J., Dentten, E., Whelan, A. O., Gordon, S. V., Kelly, G., Frenkel, T. A., Hewinson, R. G., and Williamson, R. A. (2003) J. Biol. Chem. 278, 43736–43743
45. Mussett, F. W., Frenkel, T. A., Fenney, J., Freedman, R. B., Carr, M. D., and Williamson, R. A. (1998) J. Biol. Chem. 273, 21736–21743
46. Koradi, R., Billeter, M., and Wüthrich, K. (1996) J. Mol. Graphics 14, 51–55
47. Sreerama, N., and Woody, R. W. (2000) Anal. Biochem. 287, 252–260
48. Fenney, J., Bauer, C. T., Frenkel, T. A., Birdsell, B., Carr, M. D., Roberts, G. C. K., and Arnold, J. R. P. (1991) J. Magn. Reson. 91, 607–613
49. Demarest, S. J., Deechongkit, S., Dyson, H. J., Evans, R. M., and Wright, P. E. (2004) Protein Sci. 13, 203–210
50. Dyson, J. H., and Wright, P. E. (2005) Nat. Rev. Mol. Cell Biol. 6, 197–208
51. Freedman, S. J., Sun, Z. Y., Kung, A. L., France, D. S., Wagner, G., and Eck, M. J. (2003) Nat. Struct. Biol. 10, 504–512
52. Rowan, B. G., Garrison, N., Weigel, N. L., and O’Malley, B. W. (2000) Mol. Cell. Biol. 20, 8720–8730
53. Wu, R. C., Qin, J., Yi, P., Wong, J., Tsai, S. Y., Tsai, M. J., and O’Malley, B. W. (2004) Mol. Cell 15, 937–949
54. Qin, B. Y., Liu, C., Srinath, H., Lam, S. S., Correia, J. J., Derynck, R., and Lin, K. (2005) Structure 13, 1269–1277