Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains

Jonathan M Wojciak,
Maria A Martinez-Yamout,
H Jane Dyson and Peter E Wright*

Department of Molecular Biology, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA, USA

CBP/p300 transcriptional coactivators mediate gene expression by integrating cellular signals through interactions with multiple transcription factors. To elucidate the molecular and structural basis for CBP-dependent gene expression, we determined structures of the CBP TAZ1 and TAZ2 domains in complex with the transactivation domains (TADs) of signal transducer and activator of transcription 2 (STAT2) and STAT1, respectively. Despite the topological similarity of the TAZ1 and TAZ2 domains, subtle differences in helix packing and surface grooves constitute major determinants of target selectivity. Our results suggest that TAZ1 preferentially binds long TADs capable of contacting multiple surface grooves simultaneously, whereas smaller TADs that are restricted to a single contiguous binding surface form complexes with TAZ2. Complex formation for both STAT TADs involves coupled folding and binding, driven by intermolecular hydrophobic and electrostatic interactions. Phosphorylation of S727, required for maximal transcriptional activity of STAT1, does not enhance binding to any of the CBP domains. Because the different STAT TADs recognize different regions of CBP/p300, there is a potential for multivalent binding by STAT heterodimers that could enhance the recruitment of the coactivators to promoters.

The EMBO Journal (2009) 28, 948–958. doi:10.1038/emboj.2009.30; Published online 12 February 2009

Subject Categories: chromatin & transcription; structural biology

Keywords: CREB-binding protein; NMR structure; p300; STAT1; STAT2

Introduction

Many signal-transduction pathways converge at transcriptional coactivator complexes, which function as nodes in complex protein–protein interaction networks that process cellular signals and regulate gene expression (Naar *et al*,

2001). One of the most extensively studied transcriptional coactivators, the CREB-binding protein (CBP) (Chrivia *et al*, 1993) and its paralogue p300 (Eckner *et al*, 1994), stimulate gene expression through histone acetyltransferase activity and integrate signals that promote cell growth, transformation, development and other essential cellular processes (Goodman and Smolik, 2000). Over 300 proteins have been classified as CBP/p300-binding partners (Kasper *et al*, 2006), and many of the interactions map to regions of CBP/p300 with remarkably high redundancy. These regions represent ‘gates’ where cellular signals converge and gene expression may be downregulated by competing nuclear factors or potentiated by cooperative recruitment of CBP/p300 and components of the transcriptional machinery.

CBP and p300 are modular proteins that contain globular domains separated by long intrinsically disordered regions (Dyson and Wright, 2005). The TAZ1 and TAZ2 domains of CBP (Ponting *et al*, 1996) (sometimes named CH1 and CH3, respectively) mediate interactions with the transactivation domains (TADs) of numerous transcription factors. Sequence identity between CBP and p300 in these modules is high (>95%), suggesting conserved function, structure and binding specificity. Even though the TAZ1 and TAZ2 domains share the same fold, four α -helices stabilized by three zinc atoms (De Guzman *et al*, 2000, 2005), they are highly selective in their interactions and bind different subsets of transactivation motifs.

To investigate the determinants of binding partner selectivity by the TAZ1 and TAZ2 domains of CBP, we characterized their interactions with selected TADs from the signal transducer and activator of transcription (STAT) family of proteins. The JAK–STAT pathway has a central function in the transmission of cytokine signals from the cell membrane to the nucleus and subsequent gene activation. Members of the STAT family participate in diverse biological processes and are constitutively activated in many human malignancies and diseases (O’Sullivan *et al*, 2007). STAT proteins are activated by tyrosine phosphorylation and homo- or heterodimerization through a conserved SH2 domain. The seven human STAT proteins contain a poorly conserved C-terminal TAD, which displays variable transactivational potency (Shen and Darnell, 2001) and directly interacts with CBP/p300 (Brierley and Fish, 2005).

The TADs of STAT1 and STAT2 bind selectively to the TAZ2 and TAZ1 domains of CBP/p300, respectively (Bhattacharya *et al*, 1996; Zhang *et al*, 1996). The C-terminal TAD of the p65 subunit of the NF- κ B complex also interacts with the TAZ1 domain (Gerritsen *et al*, 1997). As CBP/p300 appear to be expressed at limiting concentrations in most cells, these transcription factors may either compete for binding to a specific TAZ domain or cooperatively recruit CBP/p300 through simultaneous TAZ1 and TAZ2 binding. Both cases have been reported: the TADs of STAT2 and p65 compete for

*Corresponding author. Department of Molecular Biology, MB2, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 N Torrey Pines Road, La Jolla, CA 92037, USA.
Tel.: +1 858 784 9721; Fax: +1 858 784 9822;
E-mail: wright@scripps.edu

Received: 13 November 2008; accepted: 20 January 2009; published online: 12 February 2009

binding to the CH1 domain of p300 *in vitro* and *in vivo* (Hottiger *et al*, 1998), and the N- and C-terminal regions of CBP cooperate to mediate the STAT1/NF- κ B synergistic expression of certain inflammatory genes (Hiroi and Ohmori, 2003). In addition to the NF- κ B pathway, CBP/p300 represents a point of convergence between the JAK/STAT and other signalling pathways, including the Ras/AP-1 (Horvai *et al*, 1997), SMAD (Ghosh *et al*, 2001) and Src/Hif-1 α pathways (Gray *et al*, 2005).

To elucidate the structural basis for coordinated gene expression and CBP/p300 recruitment by STAT proteins, we have determined the structures of the TAZ1 domain in complex with the TAD of STAT2 and the TAZ2 domain in complex with the TAD of STAT1. Both TADs are intrinsically unstructured and undergo folding transitions upon binding to TAZ, leading to formation of an intermolecular hydrophobic core. Interestingly, although the binding mechanisms in the two complexes are similar and the TAZ1 and TAZ2 domains are topologically similar, the structures of the bound STAT1-TAD and STAT2-TAD are very different. The subtle differences in helix packing between TAZ1 and TAZ2 result in significantly different surface grooves for binding of targets, and appear to be the primary determinants of target selectivity.

Results

Binding affinity and selectivity

Previous studies have mapped interactions between the C-terminal TADs of STAT1 (Zhang *et al*, 1996), STAT2 (Bhattacharya *et al*, 1996), STAT3 (Schuringa *et al*, 2001) and STAT4 (O'Sullivan *et al*, 2004) to regions of CBP encompassing the TAZ1 or TAZ2 domains (Figure 1A). To investigate the relative binding affinities of these TADs for the isolated TAZ1 or TAZ2 domains, we performed *in vitro* pull-down experiments. Amino-acid sequences from the human STAT1 (residues 710–750), STAT2 (residues 769–851), STAT3 (residues 719–764) and STAT4 (residues 702–748) proteins (Figure 1B and C) were fused to the C terminus of the B1 domain from protein G (GB1). The GB1-STAT fusions were bound to IgG resin, incubated with either the isolated TAZ1 domain (residues 340–439) or TAZ2 domain (1764–1855) of CBP. All of the GB1-STAT fusion constructs interacted to some extent with TAZ2, but binding to the TAZ1 domain was more selective (Figure 2A). The GB1-STAT1 and GB1-STAT2 constructs showed higher affinity for TAZ2 compared with the GB1-STAT3 and GB1-STAT4 TADs. The TAZ1 domain showed preferential binding to the GB1-STAT2 TAD. A shorter construct of STAT2 (residues 786–838, C793H/C809S) used for structural studies retained high-affinity binding to TAZ1 but not to TAZ2, suggesting that the deleted region contains a TAZ2-binding site or enhances nonspecific binding to TAZ2.

To quantify the high-affinity interactions observed in the pull-down assay, we determined dissociation constants (K_d) using isothermal titration calorimetry (ITC). The complex formed between GB1-STAT1(710–750) and TAZ2 has an apparent K_d of 52 ± 31 nM, whereas that between GB1-STAT2(786–838, C793H/C809S) and TAZ1 has a K_d of 58 ± 3 nM. Binding to the non-cognate TAZ domains is ~ 100 -fold weaker; STAT1 binding to TAZ1 has a K_d of 4 ± 1.4 μ M and STAT2 binding to TAZ2 has a K_d of 5 μ M.

Serine phosphorylation within the STAT1, STAT3 and STAT4 TADs (Figure 1B) increases transcriptional activity

Figure 1 Characterization of STAT-TAD and TAZ domains. (A) Domain structure of mouse CBP showing the positions of the TAZ1 and TAZ2 domains in the mouse sequence. (B) Alignment of the human STAT1, STAT3 and STAT4 TAD amino-acid sequences. Bold-faced type highlights amino acids that are similar between family members. Asterisks indicate residues in the STAT1-TAD that directly contact TAZ2 in the complex structure. The grey bar denotes amino acids that form the helix in the bound STAT1-TAD. The circled 'P' denotes the position of the phosphorylation site. A dash in the sequence indicates a gap with no corresponding residue. (C) Amino-acid sequences of the human STAT2 C-terminal TAD (ctTAD) sequences and the truncated construct with two point mutations (nmrTAD). Underlined residues indicate an imperfect repeat in the STAT2-ctTAD sequence. A dash in the sequence indicates a gap with no corresponding residue. A dot in the STAT2-nmrTAD sequence indicates a position where the amino acid is identical to STAT2-ctTAD. Asterisks indicate residues in the STAT2-nmrTAD that contact TAZ1 in the complex structure. Grey bars denote the amino acids that comprise the three helical structures in the bound STAT2-nmrTAD.

and CBP/p300 recruitment *in vivo* (Wen *et al*, 1995; Decker and Kovarik, 2000). To determine the effect of phosphorylation on CBP binding, we expressed GB1-STAT1 TAD phosphorylated on S727 and measured binding affinities for TAZ1 and TAZ2 by ITC. No change in affinity was observed relative to the non-phosphorylated TAD ($K_d = 87 \pm 20$ nM and 3.7 ± 0.6 μ M for the TAZ2 and TAZ1 domains, respectively). These results were confirmed by pull-down assays that showed that phosphorylation does not change the affinity for the KIX domain, the binding site of the phosphorylated CREB TAD (Parker *et al*, 1996) and impairs binding to the NCBP (Figure 2B).

Mapping STAT TADs for structure determination

The minimal high-affinity binding site for TAZ1 on the STAT2 TAD was identified by screening the interactions of a set of truncated TAD constructs with 15 N-labelled TAZ1 using 2D 1 H- 15 N HSQC spectra. The region from G748 to the C terminus (F851) was mapped and a construct spanning residues V786-S838 found to be optimal for NMR analysis. Two mutations (C793H and C809S) were made to enhance sample stability; these cysteines are poorly conserved among

Figure 2 Interactions of TAZ and STAT domains. (A) SDS-PAGE analysis of mouse TAZ1 and TAZ2 assayed for interactions with the immobilized human STAT sequences shown in Figure 1B and C. GB1 (lanes 3 and 10), GB1-STAT1-ctTAD (lanes 4 and 11), GB1-STAT2-ctTAD (lanes 5 and 12), GB1-STAT2-nmrTAD (lanes 6 and 13), GB1-STAT3-TAD (lanes 7 and 14) and GB1-STAT4-TAD (lanes 8 and 15) bound to IgG Sepharose were incubated with purified TAZ1 (lanes 3–8) or TAZ2 (lanes 10–15). The bands at ~25 and ~50 kDa (lanes 3–8 and 10–15) correspond to IgG light and heavy chains. Lanes 2 and 9 contain free TAZ1 and TAZ2, respectively. (B) Results of pull-down assays showing that phosphorylation of S727 does not enhance the affinity of STAT1 TAD for CBP. TADs of STAT1 and MLL (used as a control for KIX binding) were bound to IgG Sepharose through an N-terminal GB1 fusion tag and binding of CBP domains TAZ1, KIX, TAZ2 and NCBD to STAT1 TAD (710–750), phosphoS727 and S727A STAT1 TADs was measured by pull-down assays monitored by reversed phase HPLC.

mammalian STAT2 proteins and the mutations result in a sequence identical to that from pig. As the full-length TAD of STAT1 is relatively short, we generated only two constructs, one starting at residue S710 and extending to residue S745 and the other from S710 to the C terminus, V750. Both peptides bind with similar affinity to TAZ2, suggesting that the C-terminal five amino acids are not essential for complex formation. Nevertheless, the longer construct (S710–V750) was used for structure determination.

The isolated STAT1-TAD and STAT2-TAD are intrinsically unstructured; the amide proton resonances occupy narrow chemical shift ranges in ^1H - ^{15}N HSQC spectra (Figure 3). Upon addition of the cognate TAZ domain, the amide proton resonances disperse and exhibit chemical shifts indicative of a stable, folded structure. Both complexes are in slow exchange on the NMR timescale. The amide proton resonances of residues 720–750 in the STAT1-TAD and residues L791–M816 of the STAT2 TAD show the largest deviations from random coil chemical shifts. The C terminus of STAT2-TAD (P817–S838) undergoes only small chemical shift perturbations and ^{15}N heteronuclear NOEs show that it remains flexible in the complex.

Figure 3 NMR spectra of free and bound STAT-TAD. (A) ^1H - ^{15}N HSQC spectrum of STAT1-TAD free (black) and bound to TAZ2 (red). (B) ^1H - ^{15}N HSQC spectrum of STAT2-TAD free (black) and bound to TAZ1 (red). Selected cross-peaks of the bound protein are labelled.

Structure determination

The structures of the STAT1-TAD:TAZ2 and STAT2-TAD:TAZ1 complexes were determined using heteronuclear multidimensional NMR. To minimize ambiguity due to spectral overlap, several samples were prepared in which one component was uniformly ^{15}N - or $^{15}\text{N}/^{13}\text{C}$ -labelled and the other component was at natural isotopic abundance. This labelling scheme enabled chemical shift assignments for >90% of the ^1H , ^{13}C and ^{15}N resonances for both components of each complex. Experimentally derived distance and dihedral angle restraints were used to calculate and refine an ensemble of 20 structures for the STAT2-TAD:TAZ1 (Figure 4A) and STAT1-TAD:TAZ2 (Figure 5A) complexes. Each TAZ domain contains four α -helices, designated H1–H4, and three zinc atoms and constitutes a pre-formed scaffold for ligand binding (De

Figure 4 NMR solution structure of the TAZ1:STAT2-TAD. (A) Ensemble of the 20 lowest energy structures superimposed on the backbone atoms of TAZ1:STAT2-TAD (TAZ1, residues Pro347-Ala372 and Cys384-Asn434; STAT2-TAD, residues Asp790-Met816). The backbones of TAZ1 and STAT2-TAD are shown in blue and green, respectively. The boundaries of the well-structured region of STAT2-TAD, and the N and C termini of TAZ1 are labelled. The backbone atoms of the flexible C-terminal region of the STAT2-TAD (residues Pro817-Ser835) are shown in light green, and the flexible loop of TAZ1 (residues 373–383) is shown in light blue. (B) Ribbon diagram showing the backbone of the lowest energy structure with zinc atoms and secondary structure elements of TAZ1 labelled. The ribbon is traced from residues Thr344-Ser436 of TAZ1 (blue) and Glu788-Met816 of the STAT2-TAD (green), with the C-terminal tail shown in light green. (C–E) Close-up views showing TAZ1 (yellow surface) and STAT2-TAD (white backbone), with the side chains of residues that form intermolecular contacts. The colour scheme for the side chains is as follows: acidic residues (Asp and Glu), red; basic residues (Arg, Lys and His), blue; neutral polar residues (Ser, Thr and Gln) and backbone carbonyl atoms, cyan; hydrophobic residues (Ala, Leu, Ile, Phe, Tyr, Val, Met and Pro), green.

Guzman *et al*, 2005). The structures in complex with the STAT2-TAD and STAT1-TAD can be superimposed on those of the respective free proteins with r.m.s. deviations of 1.8 Å for TAZ1 and 1.5 Å for TAZ2 for backbone heavy atoms (N, C α , C').

Structure of the bound STAT2-TAD

In the STAT2-TAD:TAZ1 complex, the STAT2-TAD wraps entirely around helix H1 of TAZ1, and contacts amino acids along a spiral groove formed by the packing of helices H1/H3, H1/H2 and H1/H4 (Figure 4B). The structure of STAT2-TAD is predominantly extended, with a short α helix (H793–H796) and another 3_{10} helix (P802–F806). The N-terminal region of the STAT2-TAD recognizes the surface of TAZ1 formed by helix H1 and the Zn $_3$ -binding site (Figure 4C), where the side chains of E788, D790 and D794 form intermolecular electrostatic interactions with the side chains of R350, R423 and H417. The hydrophobic groove containing the side chains of I353, I414 and V428 in TAZ1 presents a favourable surface for interactions with L791, L795 and L798 of the STAT2-TAD. The walls of this groove are formed by H417 and P427, which contact P792 and P789. The side chain of M803 also contributes to this intermolecular hydrophobic

core. Residues P802–F806 form a short helical structure, with the aromatic ring of F806 stacked between two imidazole rings (H407 and H383) and in contact with the methyl groups of L359 and M387 (Figure 4D). The side chain of R807 forms a salt bridge with the carboxylate of E348, whereas the backbone is positioned by a hydrogen bond between the carbonyl oxygen and the N ϵ of Q355. Residues N808 and S809 are solvent exposed and form a turn that positions residues V810–M816 for interactions with a groove formed by helix H1, the Zn1-binding site and the N terminus of helix H4 in TAZ1 (Figure 4E). The side chains from residues V358, L359, L361, L381, W418, P431, V432 and A435 form a large hydrophobic surface that accommodates the side chains of V810, I812, I815 and M816 of the STAT2 TAD. The C terminus (P817–S838) of the bound STAT2 is more flexible, yet residues in this region contact the TAZ1 domain. The side chains of V829 and Y833 pack against an exposed hydrophobic patch formed by the side chains of L391, M394 and I415, and K419 forms an electrostatic interaction with D830 in the majority of structures in the ensemble. These interactions are not essential for TAZ1 binding but apparently contribute to the overall stability of the complex.

Figure 5 NMR solution structure of the TAZ2:STAT1-TAD. (A) Ensemble of the 20 lowest energy structures superimposed on the backbone atoms of TAZ2:STAT1-TAD (TAZ2, residues Gln1766-Arg1852; STAT1-TAD, residues Asp721-Asn748). The backbones of TAZ2 and STAT1-TAD are shown in purple and yellow, respectively. The N and C termini of TAZ2 and STAT1-TAD are labelled. (B) Ribbon diagram showing the backbone of the lowest energy structure with zinc atoms and secondary structure elements of TAZ2 labelled. (C-E) Close-up views showing TAZ2 (yellow surface) and STAT1-TAD (white backbone), with the side chains of residues that form intermolecular contacts. The colour scheme for the side chains is the same as for Figure 4. The position of the side chain of S727, the phosphorylation site characterized in Figure 1, is shown in magenta.

Structure of the bound STAT1-TAD

The structure of STAT1-TAD bound to TAZ2 differs dramatically from that of STAT2-TAD bound to TAZ1. Unlike the extended surface over which the STAT2-TAD binds TAZ1, the STAT1-TAD is localized to one face of TAZ2 and adopts a single well-defined helix (residues P728-V738) that is longer than any of the STAT2 helices. This helix binds in a large hydrophobic surface created by the packing of helices H1, H2 and H3 of TAZ2 (Figure 5B). Residues D721-S740 of the STAT1-TAD form a wedge, with its vertex at P725 that penetrates into a deep hydrophobic cavity formed by the intersection of helices H1, H2 and H3 and packs against the side chains of M1799, V1802, L1826 and the aliphatic portion of Q1822 (Figure 5C). The side chain of Q1822 forms a buried hydrogen bond with the backbone carbonyl group of M726 and the methionine side chain projects deeply into the TAZ2 domain to contact the methyl groups of A1825 and L1826. The N-terminal region of the STAT1-TAD adopts an extended backbone conformation that is stabilized by hydrophobic contacts from L724 to P1818 and V1819 and an intermolecular salt bridge between the side chains of D721 and R1811. Within the amphipathic α -helix between residues P728-V738, the side chains of F731, V734, I737 and V738 form a hydrophobic face that contacts the side chains of I1773, A1825, L1826 and Y1829 located in helices H1 and H3 (Figure 5D).

The opposite face of the helix is negatively charged and consists of E729, E730, D732 and E733, which form electrostatic interactions with the guanidinium groups of R1769, R1770 and R1775 from helix H1 of TAZ2. The helix terminates at G739, which forms a turn and positions the C-terminal portion of the STAT1-TAD to rest in the cleft between the Zn₂-binding site, helix H3 and helix H4 (Figure 5E). A hydrophobic patch formed by the side chains of T1813, L1824, I1847, L1851 and F1843 and the aliphatic portions of K1821 and K1850 are contacted by the side chains of V741, F743, M746 and M747. Residues at the N terminus of STAT1 (S710-D721) are disordered and do not contact TAZ2 in the structures.

Both TADs bury considerable surface area upon complex formation, $\sim 2300 \text{ \AA}^2$ for TAZ1:STAT2-TAD and $\sim 1600 \text{ \AA}^2$ for TAZ2:STAT1-TAD. These extensive interfaces encompass numerous highly specific intermolecular interactions that constitute the molecular basis for high-affinity binding and selective recruitment of the CBP TAZ2 and TAZ1 domains, respectively, by the STAT1 and STAT2 TADs.

Complementary electrostatics mediate recognition of acidic TADs

The surfaces of both TAZ domains are highly electropositive (Figure 6), favouring binding of acidic TADs. The electrostatic

Figure 6 Electrostatic potentials on the surfaces of the TAZ domains. Positively charged surfaces are shown in blue, and negatively charged in red. (A) TAZ2 in complex with the STAT1-TAD (yellow backbone). (B) TAZ1 in complex with STAT2-TAD (green backbone). The left and right images represent a 180° rotation around the vertical axis in the plane of the page. (C) Superposition of the structure of STAT2-TAD (green) on the TAZ2:STAT1-TAD complex (transparent surface plus yellow backbone). (D) Superposition of the structure of STAT1-TAD (yellow) on the TAZ1:STAT2-TAD complex (transparent surface plus green backbone).

interactions between the TAZ2 domain and the STAT1-TAD are localized to one face of the TAZ2 scaffold (Figure 6A) and arise exclusively from acidic residues in STAT1-TAD contacting basic residues in TAZ2. The opposite surface of TAZ2 is composed of mostly uncharged amino acids and appears less competent for binding negatively charged TADs. By contrast, the surface of the TAZ1 domain has a more even distribution of positive charges (Figure 6B). The STAT2-TAD wraps around the TAZ1 domain, forming multiple electrostatic interactions with surface positive charges. In addition, the positively charged side chain of R807 in STAT2-TAD forms an electrostatic interaction with the negatively charged side chain of E348 of TAZ1. The preferential binding of acidic activation domains by TAZ domains and the evidence for charge complementarity of the bound STAT1-TAD and STAT2-TAD suggest that electrostatic interactions have a key function in determining the specificity of TAD recognition, and the intermolecular hydrophobic core that forms upon TAD folding provides stability to the complex.

Structural basis for TAD recognition

Aside from specific intermolecular side chain interactions, which predicate the precision of the sequence specificity, the

structural differences between TAZ1 and TAZ2 provide a further basis for binding partner recognition and discrimination. Comparison of the free TAZ1 and TAZ2 structures reveals that, even though the overall topologies are very similar, there are significant differences in the orientation of helix H4 (De Guzman *et al*, 2000, 2005). Helix H4 of TAZ1 forms part of the binding surface for the C-terminal activation domains of HIF-1 α and CITED2 (De Guzman *et al*, 2005) and appears also to be a major determinant for recognition of the STAT2-TAD. The C terminus of helix H4 packs against helix H1 in TAZ1, forming a hydrophobic pocket that binds residues I812-M816 of the STAT2-TAD. Both surfaces are dramatically different when helix H4 is positioned as in the TAZ2 structure (Figure 6C and D).

In addition to differences in the orientation of helix H4, differences in the packing of helix H3 also appear to have a function in binding partner discrimination by TAZ domains. The helix H2-H3 packing angle differs by $\sim 35^\circ$ in the structures of TAZ1 and TAZ2 (De Guzman *et al*, 2000, 2005). The position of H3 in TAZ2 creates an exposed hydrophobic pocket contacted by the LPMSF motif in STAT1-TAD (Figure 6C). In TAZ1, the change in the packing angle of helices H3 and H1 occludes this binding pocket (Figure 6D). This is consistent with the binding assays, which show only very weak interaction between TAZ1 and the STAT1, STAT3 and STAT4 TADs containing the LPMSF motif (Figure 2A). A model of the STAT2-TAD superimposed onto the structure of TAZ2 (Figure 6C) predicts a more favourable interaction devoid of such steric clashes, and this too is consistent with the ability of GB1-STAT2 to pull-down TAZ2 *in vitro*.

The helix packing in TAZ2 creates a contiguous U-shaped interface for protein binding. The centre of this interface contains a hydrophobic groove that contacts the helix (residues P728-V738) in the bound STAT1-TAD structure (Figure 6A). The groove is formed by the side chains of R1775 and K1832, which are ~ 20 Å apart in the structure, and the side chains of R1769 and K1821, which are ~ 18 Å apart. The structure and charge distribution of this groove present a favourable surface to bind a negatively charged, amphipathic helix. The groove is wide, suggesting it could accommodate a helix with bulky side chains (such as F743 in STAT1-TAD) or slightly different orientations without causing steric clashes. Adjacent to the helix-binding groove is a deep and narrow hydrophobic cleft that contacts the LPMSF motif in the STAT1-TAD:TAZ2 complex and favours binding of extended structure rather than helix. The combination of these features presents a contiguous surface that binds and stabilizes the compact structure of the bound STAT1-TAD.

The helices in TAZ1 pack to form a spiral groove in the surface that supports coupled folding and binding of the STAT2-TAD. Although the dimensions of this groove vary, it is in general narrower than the helix-binding groove and shallower than the LPMSF-binding cleft of TAZ2. Unlike the relatively flat protein-binding surface on TAZ2, the groove in TAZ1 wraps around the structure presenting an extensive concave surface for protein binding. This feature induces the STAT2-TAD to adopt an extended structure, with three short helical turns that contact isolated regions along the spiral groove.

Discussion

The structures of the STAT1 and STAT2 complexes reported here reveal different modes of binding of transcriptional activation domains by the TAZ1 and TAZ2 domains of CBP/p300. The distinct topologies of the TAZ1 and TAZ2 domains appear to form the basis for binding partner selectivity (De Guzman *et al*, 2005). The surface of the TAZ1 domain contains interlinked hydrophobic grooves that favour interactions with long regions of intrinsically disordered TADs that can form one or more local helical segments upon binding and also interact with the narrower grooves of TAZ1 in relatively extended conformations. Many of these grooves are absent from the surface of the TAZ2 domain. The relative affinities with which an intrinsically disordered TAD binds to TAZ1 or TAZ2 depend upon its ability to fold and form the appropriate stabilizing interactions, as discussed in detail below.

TAZ recognition by other binding partners

Structures of the TAZ1 domain have been solved in complex with a number of different partners, allowing us to compare the various modes of binding by TADs with little sequence homology (Dames *et al*, 2002; Freedman *et al*, 2002, 2003; De Guzman *et al*, 2004). The structures of the bound C-terminal TADs of HIF-1 α , CITED2 and STAT2 occupy the same spiral groove on TAZ1 (Figure 7A). The most striking difference between these TADs is their orientation on TAZ1; looking down the long axis of helix H1 from the N to C terminus, the TADs of HIF-1 α and CITED2 wrap around TAZ1 in a clockwise direction (N to C terminus), where STAT2-TAD is bound in the opposite (counterclockwise) orientation. The area of TAZ1 bound by STAT2-TAD more closely resembles that bound by CITED2-TAD than by HIF-1 α -TAD, which forms a stable helix at the C terminus that contacts an additional surface on TAZ1 (Figure 7A). Residues V227, L228 and L231 in the N-terminal helix of CITED2 make contacts similar (but in the opposite direction) to those of residues M816, I815 and I812 in STAT2-TAD (Figure 7B). This groove on TAZ1 appears well suited to support folding and binding of an amphipathic helix of varying length and directionality. The HIF-1 α and CITED2 TADs contain a conserved LPXL motif, which forms an extended structure and contacts hydrophobic side chains within the helix H1–H3 groove of TAZ1. By contrast, residues P802–F806, which contact this region of TAZ1, form a helical structure in the STAT2-TAD:TAZ1 complex (Figure 7C). The lack of directionality among bound partners, along with the presence of different secondary structures contacting the same regions of TAZ1, demonstrates the promiscuous nature of TAZ1 recognition and is consistent with the apparent lack of a full consensus binding sequence.

The STAT1-TAD:TAZ2 structure is the first reported structure of the TAZ2 domain in complex with a binding partner. Chemical shift perturbation studies indicated that a region of the TAD of the p53 tumour suppressor binds to the same surface of TAZ2 as the STAT1-TAD (De Guzman *et al*, 2000). The binding mode of the STAT1-TAD in the TAZ2 structure may therefore represent a common mode of binding, which may allow identification of other sequences that are likely TAZ2-binding partners.

Residues P728–V738 in STAT1-TAD form extensive intermolecular electrostatic and hydrophobic contacts and under-

Figure 7 Comparison of TAZ1 ligand structures. (A) Surface representation of the structure of TAZ1 in complex with the STAT2-TAD (green), HIF-1 α -CTAD (red) and CITED2-TAD (blue). The left and right images represent a 180° rotation around the vertical axis in the plane of the page. The N and C termini of each ligand are labelled. (B) Superposition of the complexes of TAZ1 with STAT2-TAD (green) and CITED2-TAD (blue), showing a region of similar structure in the two complexes, even though the sequences run in opposite directions. (C) Superposition of the complexes of STAT2-TAD (green) and CITED2-TAD (blue), showing different secondary structures bound to the same region of TAZ1. (D) Sequence alignment showing proteins with homology to the helical region of STAT1-TAD. Amino acids are coloured red (acidic), blue (basic), yellow (non-charged polar), green (hydrophobic) and purple (Pro and Gly).

go the greatest degree of structural ordering upon binding TAZ2. Therefore, we used this sequence of STAT1-TAD as the basis for pattern recognition searches of protein sequences in the Swiss-Prot database. Several known TAZ2-binding proteins (SV40 large T antigen, p53 and members of the GATA family of transcriptional activators) contain sequence motifs that are similar to the STAT1-TAD (Figure 7D), particularly in the possession of a sequence pattern P-D/E-D/E- Φ -(D/E or Φ)-(D/E or Φ), where D/E refers to Asp or Glu and Φ is a hydrophobic amino acid. Residues L43–F54 within the p53 TAD exhibit strong homology to the LPMSMP motif and residues in the helix of STAT1-TAD, and residues within this region of the p53 TAD are critical for TAZ2 binding (Teufel *et al*, 2007). Members of the MAF family of transcriptional activators also have regions with high homology to the STAT1-TAD. Our search identified several motifs in viral proteins with high sequence similarity to STAT1-TAD; it is inviting to speculate that inactivation of CBP through direct competition for the TAZ2 domain may be a common mechan-

ism for viral hijacking of host transcriptional machinery. The most similar sequence retrieved from this search belongs to the NEF protein from the HIV-1 MAL isolate. Interestingly, the region in NEF (residues 147–159) that is highly homologous to the STAT1-TAD is located at the beginning of a large, internal, flexible loop (Grzesiek *et al*, 1996; Lee *et al*, 1996), raising the possibility that this region may be capable of coupled folding and binding to cellular proteins, such as the TAZ2 domain of CBP/p300. Although NEF has been implicated in the regulation of cellular gene expression (Pessler and Cron, 2004), to our knowledge no direct interaction with CBP/p300 has been reported.

CBP/p300 recruitment by the STAT family of proteins

All seven members of the human STAT family have been reported to interact with CBP/p300 through their C-terminal TADs (Bhattacharya *et al*, 1996; Zhang *et al*, 1996; Pfizner *et al*, 1998; Schuringa *et al*, 2001; O'Sullivan *et al*, 2004, 2007). The LPMSP motif, which is conserved within the TADs of STAT1, STAT3 and STAT4, mediates the recruitment of CBP/p300 by STAT1 (Sun *et al*, 2005) and STAT3 (Sun *et al*, 2006). In the STAT1-TAD:TAZ2 complex, residues L724, P725 and M726 of this motif form intimate contacts with TAZ2 through hydrophobic interactions from the side chains and hydrogen bonds involving the backbone carbonyl oxygen atoms (Figure 5C). Although the LPMSP motif forms several important contacts with TAZ2, these are not sufficient for high-affinity binding. Residues following the LPMSP motif, which are not conserved between STAT1, STAT3 and STAT4, appear to be the major determinant for high-affinity TAZ2 binding. In the STAT1-TAD:TAZ2 complex, the side chains of E729, E730, D732 and E733 interact with arginine side chains on the surface of TAZ2. In the STAT3-TAD and STAT4-TAD, many residues at these positions are not negatively charged and incapable of forming such intermolecular interactions. This may explain the stronger TAZ2 binding observed *in vitro* for the STAT1-TAD compared with STAT3-TAD and STAT4-TAD (Figure 2A).

Previous studies have indicated that phosphorylation of S727 within the LPMSP motif is required for maximal transcriptional activity of STAT1 and STAT3 (Wen *et al*, 1995; Decker and Kovarik, 2000) and it has been suggested that this may reflect enhanced binding of the phosphorylated TAD to CBP/p300 (Zakharova *et al*, 2003). A S727A mutation retains DNA-binding activity, but histone acetylation and recruitment of CBP are impaired *in vivo* (Varinou *et al*, 2003). In this study, we show that S727 phosphorylation in the STAT1 TAD causes no change in binding affinity for the TAZ2, TAZ1 or the KIX domain of CBP, whereas both phosphorylation and an S727A mutation appear to impair binding to the nuclear coactivator-binding domain (residues 2059–2117) of CBP. The S727 side chain, which caps the N terminus of the STAT1 helix, points towards the solvent (Figure 5C) and can readily accommodate a phosphate group without perturbing the intermolecular interactions. Thus, our results suggest that S727 phosphorylation does not enhance binding to CBP/p300 but stimulates the activity of STAT1 and STAT3 by altering interactions with other components of the transcriptional machinery. Indeed, it has recently been shown that S727 phosphorylation occurs only after STAT1 has been assembled into a chromatin-associated transcription complex (Sadzak *et al*, 2008), and that it functions by regulating STAT1

SUMOylation (Vanhatupa *et al*, 2008). Phosphorylation of the STAT1 TAD has also been shown to enhance transcriptional activity by accelerating nucleocytoplasmic shuttling (Lodige *et al*, 2005). It appears that phosphorylation of serine residues in the C-terminal TAD is not a requirement for the recruitment of the CBP/p300 coactivators by STAT1 (and presumably also STAT3 and STAT4).

The STAT family of proteins has evolved to target different domains in CBP/p300; differences in the sequences of the C-terminal TADs enable the STAT proteins to recruit CBP/p300 by targeting different domains of the transcriptional coactivators. Thus, the STAT1 TAD binds TAZ2 and the NCBD domains (Figure 2A and B), the STAT2 TAD binds TAZ1 (and to a lesser extent TAZ2—Figure 2A), the STAT3 and STAT4 TADs bind TAZ2, whereas the STAT5 and STAT6 TADs recognize the KIX and NCBD domains, respectively (Pfizner *et al*, 1998; Gingras *et al*, 1999). Thus, STAT heterodimers can potentially recruit transcriptional coactivators with enhanced efficiency by simultaneously targeting two domains of CBP/p300; for example, STAT2:STAT1 and STAT2:STAT3 heterodimers can bind CBP/p300 through both the TAZ1 and TAZ2 domains. The STAT1 homodimer also has the potential for multivalent interactions with the TAZ2 and NCBD regions of CBP/p300. In addition, activated STAT dimers have been shown to form higher order complexes on promoters containing multiple recognition sites (Vinkemeier *et al*, 1996; Xu *et al*, 1996; John *et al*, 1999), thereby further increasing the number and complexity of STAT TADs available to mediate interactions with coactivators and the general transcriptional machinery.

Materials and methods

Protein expression and purification

Proteins were expressed from a coexpression plasmid (Sugase *et al*, 2008), transformed into *Escherichia coli* BL21 (DE3) (+ DNAY) cells. Human STAT1 (residues 710–750) or STAT2 (residues 786–838, C793H/C809S) were cloned immediately downstream of a His₆-GB1 tag to generate a fusion construct with a thrombin site between the GB1 and STAT-TAD. TAZ1 (mouse CBP residues 340–439) or TAZ2 (mouse CBP residues 1764–1855) genes were inserted into the second coding region. Cells were grown at 37°C in M9 media, supplemented with ¹⁵NH₄Cl or (¹⁵NH₄)₂SO₄ and ¹³C₆-glucose, to OD₆₀₀ of 0.6–0.8. After induction with 1 mM IPTG, the medium was supplemented with 150 µM ZnSO₄ and growth continued at 15°C for 12–16 h.

Cells were sonicated in lysis buffer (25 mM Tris (pH 8), 250 mM NaCl, 25 mM imidazole and 6 M urea). Ni-NTA superflow resin (Qiagen) was added to the soluble fraction, incubated for 30 min and then poured into a gravity flow column. The flow-through containing TAZ1 was loaded onto SP-Sepharose equilibrated in 25 mM Tris (pH 8) and 6 M urea, and TAZ1 was eluted with a linear gradient to 1 M NaCl. The Ni-NTA resin containing bound His₆-GB1-STAT-TAD was washed with lysis buffer (without urea), incubated with thrombin for ~4 h, and the cleaved STAT-TAD was collected from the supernatant. The STAT-TAD cleavage products contain four additional N-terminal residues (GSHM) from the vector. The STAT-TAD and TAZ1 fractions were further purified by reversed phase HPLC in acetonitrile/TFA buffer. Cells containing a His₆-GB1-STAT1-TAD/TAZ2 coexpression plasmid were lysed by sonication in a buffer containing 20 mM Tris, 8 M urea, 200 mM arginine hydrochloride, 40 mM NaCl, 10 mM DTT, pH 8. The soluble protein was applied to SP-Sepharose equilibrated in lysis buffer and TAZ2 was eluted with a linear gradient to 500 mM NaCl. The protein was further purified by reversed phase HPLC.

STAT1-TAD was phosphorylated at S727 using activated p38α kinase. GST-MKK6 was expressed in Rosetta DE3 pLysS cells with overnight induction at 20°C. Cells were lysed in 20 mM Tris pH 8,

100 mM NaCl, 5 mM DTT, 1 mM EDTA, 10% glycerol and 0.1% Triton X-100, and soluble protein was affinity purified on GST Sepharose 4B. His₆-p38 α was expressed in Rosetta DE3 pLysS with overnight induction at 15°C. Cells were lysed in 20 mM Tris pH 8, 250 mM NaCl, 10 mM imidazole, 5 mM β -mercaptoethanol and 0.2% Triton X-100, and soluble protein was purified by metal affinity chromatography on NiNTA resin. Purified GST-MKK6, His₆-p38 α and His₆GB1-STAT1 TAD were exchanged into 20 mM HEPES pH 7.5, 50 mM NaCl. Kinase reactions utilized 140 μ M His₆GB1-STAT1 TAD, 15 μ M His₆-p38 α , 2 μ M GST-MKK6 in 50 mM HEPES pH 7.5, 0.1 mM PMSF, 1 mM NaF, 0.1 mM sodium orthovanadate, 0.1 mM β -glycerol phosphate, 10 mM MgCl₂, 5 mM DTT and 10 mM disodium ATP and were performed at room temperature for 24 h. Phosphorylated His₆GB1-STAT1 TAD was purified by reversed phase HPLC as described above. MALDI analysis of a tryptic digest showed a mass increase corresponding to a single phosphate. In control reactions, the S727D STAT1 mutant remained unphosphorylated, confirming that p38 α phosphorylation is specific to S727.

GB1 pull-down assay

Purified GB1 or GB1-STAT fusion protein was added to an 8% (w/v) slurry of IgG-Sepharose 6 Fast Flow (GE Healthcare Biosciences) in assay buffer (20 mM Tris-HCl (pH 8), 150 mM NaCl, 0.1% (v/v) NP-40) and washed three times with assay buffer. Equimolar TAZ1 or TAZ2 was added to the resin-bound GB1 or GB1-STAT and incubated for 15 min at room temperature. Unbound proteins were removed by washing two times with assay buffer and bound proteins were analysed using SDS-PAGE and stained with Coomassie brilliant blue.

Binding of the STAT1 TAD, the S727A mutant, and the phosphoS727 TAD to the TAZ1, TAZ2, NCBP and KIX domains of CBP was assayed by GB1 pull-down experiments monitored by reversed phase HPLC. GB1-MLL (2842–2869), GB1-p53 (1–94) CBP KIX (586–672) and NCBP (2059–2117) were purified as described previously (Demarest *et al*, 2002; Goto *et al*, 2002; Legge *et al*, 2004). His₆GB1 produced by thrombin cleavage of His₆GB1-p53 was used as a control for nonspecific binding. All proteins for pull-down experiments were exchanged into 20 mM Tris pH 8, 50 mM NaCl and 1 mM DTT. For each pull-down assay, 30 ml of IgG-

Sepharose 6 Fast Flow (GE Healthcare Biosciences) was equilibrated in assay buffer (20 mM Tris pH 8, 50 mM NaCl and 0.1% IGEAL CA-630) and loaded with 6 nmol GB1-TAD. Resin-bound GB1-TADs were incubated with 6 nmol CBP domains (TAZ1, TAZ2, KIX or NCBP) for 15 min at room temperature. Unbound proteins were removed by washing with assay buffer. Bound proteins were eluted from the resin by the addition of 0.1% trifluoroacetic acid and monitored by reversed phase HPLC chromatography on a Jupiter 5 mm C4 analytical column (Phenomenex). Results were corrected for nonspecific binding of CBP domains by subtracting HPLC peak areas derived from a second set of control experiments with His₆GB1 in place of His₆GB1-TAD. Representative chromatograms are shown in the Supplementary data.

ITC

TAZ1 or TAZ2 were titrated with 10-fold excess of GB1-STAT2 and GB1-STAT1 in ITC buffer (20 mM Tris-HCl (pH 8), 50 mM NaCl and 1 mM DTT) at 35°C using a VP-ITC MicroCalorimeter (MicroCal). One injection of 5 μ l was followed by 29 injections of 10 μ l until a molar ratio of $\sim$ 3.0 GB1-STAT/TAZ was obtained. The binding isotherm was fitted to a one-site binding model using the Origin 7.0 software (MicroCal). The stoichiometric ratio, the dissociation constant (K_d) and the change in enthalpy were determined using a nonlinear least-squares fitting algorithm. No binding was detected between isolated GB1 and TAZ1 or TAZ2. Representative binding isotherms for all of the complexes are shown in the Supplementary data.

NMR sample preparation

TAZ1 and TAZ2 were refolded in the presence of Zn²⁺ as described previously (De Guzman *et al*, 2000, 2005). The lyophilized STAT TADs were dissolved in 10 mM Tris, 50 mM NaCl, 5 mM DTT, pH 7 and added to refolded TAZ solutions, which were concentrated and dialysed against NMR buffer (10 mM d₁₁ Tris, 50 mM NaCl and 7 mM d₁₀ DTT, pH 6.8). NMR samples were $\sim$ 0.8–1.1 mM in labelled protein and contained a 20% excess of the unlabelled component. Six NMR samples were prepared for each complex: ¹⁵N TAZ:unlabelled STAT-TAD in H₂O, ¹⁵N STAT-TAD:unlabelled TAZ in H₂O, ¹³C,¹⁵N TAZ:unlabelled STAT-TAD in H₂O, ¹³C,¹⁵N STAT-TAD:unlabelled TAZ in H₂O, ¹³C,¹⁵N TAZ:unlabelled STAT-TAD in D₂O and ¹³C,¹⁵N STAT-TAD:unlabelled TAZ in D₂O.

Table 1 NMR and refinement statistics for protein structures

	TAZ1:STAT2-TAD	TAZ2:STAT1-TAD
<i>NMR distance and dihedral constraints</i>		
Distance constraints		
Total NOE	1711	1507
Intra-residue	168	200
Inter-residue		
Sequential ($ i-j = 1$)	417	431
Medium range ($ i-j < 4$)	365	396
Long range ($ i-j > 5$)	202	193
Intermolecular	559	287
Hydrogen bonds	0	0
Total dihedral angle restraints		
ϕ	78	72
ψ	76	72
χ_1	51	48
<i>Structure statistics</i>		
Violations (mean and s.d.)		
Distance constraints (Å)	0.084 $\pm$ 0.004	0.087 $\pm$ 0.004
Dihedral angle constraints (deg)	0	0
Maximum dihedral angle violation (deg)	0	0
Maximum distance constraint violation (Å)	0.27	0.36
Deviations from idealized geometry		
Bond lengths (Å)	0.01094 $\pm$ 0.00009	0.01069 $\pm$ 0.00006
Bond angles (deg)	2.79 $\pm$ 0.03	2.63 $\pm$ 0.03
Average pairwise r.m.s. deviation ^a (Å)		
Heavy	1.30 $\pm$ 0.14	1.10 $\pm$ 0.12
Backbone	0.61 $\pm$ 0.16	0.46 $\pm$ 0.10

^aPairwise r.m.s. deviation was calculated among 20 refined structures for the regions (TAZ1): Pro347-Ala372, Cys384-Asn434; (STAT2-TAD): Asp790-Met816; (TAZ2) Gln1766-Arg1852; (STAT1-TAD): Asp721-Asn748.

NMR spectroscopy and structure calculations

NMR spectra were acquired at 290 K (TAZ2:STAT1-TAD complex) or 305 K (TAZ1:STAT2-TAD complex) on Bruker DRX600 MHz and AVANCE 900 MHz spectrometers, processed with NMRPipe (Delaglio *et al*, 1995), and analysed using NMRView (Johnson and Blevins, 1994). Backbone resonances were assigned using HNCACB (Wittekind and Mueller, 1993), CBCA(CO)NH (Grzesiek and Bax, 1992a), HNCA, HN(CO)CA and HNCO (Grzesiek and Bax, 1992b) experiments. Assignments of side chain resonances were made using 3D HCCH-COSY and HCCH-TOCSY experiments (Bax *et al*, 1990). Stereospecific assignments and coupling constants were obtained from 3D HACAHB-COSY (Grzesiek *et al*, 1995), HNHB (Archer *et al*, 1991) and HNHA (Vuister and Bax, 1993) spectra. Distance restraints were derived from 3D ^{15}N -edited NOESY-HSQC ($\tau_m = 120$ ms), ^{13}C -edited NOESY-HSQC ($\tau_m = 100$ ms) and ^{13}C -edited, ^{12}C -filtered NOESY-HSQC ($\tau_m = 200$ ms) spectra.

Distance restraints and initial structures were generated with the programs CANDID (Herrmann *et al*, 2002) and CYANA (Güntert, 2004). As indicated by coupling constants, NOE patterns and the program TALOS (Cornilescu *et al*, 1999), backbone dihedral angles for residues in helical regions were restrained to $-60^\circ \pm 30^\circ$ (ϕ) and $-40^\circ \pm 30^\circ$ (ψ), and several side chain χ_1 angles were restrained to $(-60^\circ, 60^\circ, 180^\circ) \pm 30^\circ$. Distance restraints to define the zinc sites and helical hydrogen bonds were imposed to facilitate automated NOE assignment but were omitted during refinement. The experimental restraints are summarized in Table I.

A total of 200 structures generated in CYANA with redundant dihedral angle constraints (Güntert and Wüthrich, 1991) were further refined through two cycles of restrained molecular dynamics-simulated annealing using the AMBER 8 software package (Case *et al*, 2005). These structures were initially subjected to 2000 steps of energy minimization, followed by 20 ps of simulated annealing in vacuum and another 20 ps of simulated annealing using a generalized Born solvent model (Tsui and Case, 2000) and finally another 2000 steps of energy minimization.

References

- Archer SJ, Ikura M, Torchia DA, Bax A (1991) An alternative 3D NMR technique for correlating backbone ^{15}N with side chain ^1H resonances in larger proteins. *J Magn Reson* **95**: 636–641
- Bax A, Clore GM, Gronenborn AM (1990) ^1H - ^1H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C -enriched proteins. *J Magn Reson* **88**: 425–431
- Bhattacharya S, Eckner R, Grossman S, Oldread E, Arany Z, D'Andrea A, Livingston DM (1996) Cooperation of Stat2 and p300/CBP in signalling induced by interferon- α . *Nature* **383**: 344–347
- Brierley MM, Fish EN (2005) Stats: multifaceted regulators of transcription. *J Interferon Cytokine Res* **25**: 733–744
- Case DA, Cheatham III TE, Darden T, Gohlke H, Luo R, Merz KMJ, Onufriev A, Simmerling C, Wang B, Woods R (2005) The Amber biomolecular simulation programs. *J Comput Chem* **26**: 1668–1688
- Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy M, Goodman RH (1993) Phosphorylated CREB binds specifically to nuclear protein CBP. *Nature* **365**: 855–859
- Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* **13**: 289–302
- Dames SA, Martinez-Yamout M, De Guzman RN, Dyson HJ, Wright PE (2002) Structural basis for Hif-1 α /CBP recognition in the cellular hypoxic response. *Proc Natl Acad Sci USA* **99**: 5271–5276
- De Guzman RN, Liu HY, Martinez-Yamout M, Dyson HJ, Wright PE (2000) Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. *J Mol Biol* **303**: 243–253
- De Guzman RN, Martinez-Yamout M, Dyson HJ, Wright PE (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
- De Guzman RN, Wojciak JM, Martinez-Yamout MA, Dyson HJ, Wright PE (2005) CBP/p300 TAZ1 domain forms a structured scaffold for ligand binding. *Biochemistry* **44**: 490–497
- Decker T, Kovarik P (2000) Serine phosphorylation of STATs. *Oncogene* **19**: 2628–2637
- Delaglio F, Grzesiek S, Vuister GW, Guang Z, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**: 277–293
- Demarest SJ, Martinez-Yamout M, Chung J, Chen H, Xu W, Dyson HJ, Evans RM, Wright PE (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**: 549–553
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**: 197–208
- Eckner R, Ewen ME, Newsome D, Gerdes M, DeCaprio JA, Lawrence JB, Livingston DM (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 Dev* **8**: 869–884
- Freedman SJ, Sun ZY, Kung AL, France DS, Wagner G, Eck MJ (2003) Structural basis for negative regulation of hypoxia-inducible factor-1 α by CITED2. *Nat Struct Biol* **10**: 504–512
- Freedman SJ, Sun ZY, Poy F, Kung AL, Livingston DM, Wagner G, Eck MJ (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 α . *Proc Natl Acad Sci USA* **99**: 5367–5372
- Gerritsen ME, Williams AJ, Neish AS, Moore S, Shi Y, Collins T (1997) CREB-binding protein p300 are transcriptional coactivators of p65. *Proc Natl Acad Sci USA* **94**: 2927–2932
- Ghosh AK, Yuan W, Mori Y, Chen SJ, Varga J (2001) Antagonistic regulation of type I collagen gene expression by interferon- γ and transforming growth factor- β . Integration at the level of p300/CBP coactivators. *J Biol Chem* **276**: 11041–11048
- Gingras S, Simard J, Groner B, Pfitzner E (1999) P300/CBP is required for transcriptional induction by interleukin-4 and interacts with Stat6. *Nucleic Acids Res* **27**: 2722–2729
- Goodman RH, Smolik S (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev* **14**: 1553–1577
- Goto NK, Zor T, Martinez-Yamout M, Dyson HJ, Wright PE (2002) Cooperativity in transcription factor binding to the coactivator

During simulated annealing, the system was heated to 1000 K for the first 2 ps, followed by 4 ps at constant temperature, and final cooling to 0 K for the remaining 14 ps. Force constants were 30 kcal mol $^{-1}$ Å $^{-2}$ for NOE restraints and 1000 kcal mol $^{-1}$ rad $^{-2}$ for dihedral angle restraints.

The 20 structures with the lowest AMBER energy were analysed using PROCHECK-NMR (Laskowski *et al*, 1996). For the TAZ1-STAT2-TAD structure, 88% of backbone dihedral angles were in the most favoured region of the Ramachandran plot, with 11.0% in additionally allowed, 0.1% in generously allowed and 0.0% in disallowed regions. The corresponding statistics for the TAZ2-STAT1-TAD structure are 88.5% most favoured, 10.5% additionally allowed, 1.0% generously allowed and 0.0% disallowed. Figures were prepared using MOLMOL (Koradi *et al*, 1996). Coordinates and restraints have been deposited with the following acquisition numbers: RCSB ID 2ka4, 2ka6; BMRB ID 16014, 16015.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank John Chung and Gerard Kroon for assistance with NMR data acquisition and spectrometer maintenance, Brian Lee and Roberto De Guzman for useful discussions and help with structure determination, and Mindy Landes, Linda Tennant, Jeff Hosea and Bryan Clarkson for valuable technical support. GST-MKK6 and His $_6$ -p38 α expression plasmids were gifts from Peiqing Sun. This study was supported by grant CA96865 from the National Institutes of Health and the Skaggs Institute for Chemical Biology. JMW was supported by a postdoctoral fellowship from the American Cancer Society (PF-04-039-01-GMC).

- CREB-binding protein (CBP). The mixed lineage leukemia protein (MLL) activation domain binds to an allosteric site on the Kix domain. *J Biol Chem* **277**: 43168–43174
- Gray MJ, Zhang J, Ellis LM, Semenza GL, Evans DB, Watowich SS, Gallick GE (2005) HIF-1 α , STAT3, CBP/p300 and Ref-1/APE are components of a transcriptional complex that regulates Src-dependent hypoxia-induced expression of VEGF in pancreatic and prostate carcinomas. *Oncogene* **24**: 3110–3120
- Grzesiek S, Bax A (1992a) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc* **114**: 6291–6293
- Grzesiek S, Bax A (1992b) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* **96**: 432–440
- Grzesiek S, Bax A, Clore GM, Gronenborn AM, Hu JS, Kaufman J, Palmer I, Stahl SJ, Wingfield PT (1996) The solution structure of HIV-1 Nef reveals an unexpected fold and permits delineation of the binding surface for the SH3 domain of Hck tyrosine protein kinase. *Nat Struct Biol* **3**: 340–345
- Grzesiek S, Kuboniwa H, Hinck AP, Bax A (1995) Multiple-quantum line narrowing for measurement of H α -H β J coupling in isotopically enriched proteins. *J Am Chem Soc* **117**: 5312–5315
- Güntert P (2004) Automated protein structure calculation with CYANA. *Methods Mol Biol* **278**: 353–378
- Güntert P, Wüthrich K (1991) Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. *J Biomol NMR* **1**: 447–456
- Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319**: 209–227
- Hiroi M, Ohmori Y (2003) The transcriptional coactivator CREB-binding protein cooperates with STAT1 and NF- κ B for synergistic transcriptional activation of the CXC ligand 9/monokine induced by interferon- γ gene. *J Biol Chem* **278**: 651–660
- Horvai AE, Xu L, Korzus E, Brard G, Kalafus D, Mullen TM, Rose DW, Rosenfeld MG, Glass CK (1997) Nuclear integration of JAK/STAT and Ras/AP-1 signaling by CBP and p300. *Proc Natl Acad Sci USA* **94**: 1074–1079
- Hottiger MO, Felzien LK, Nabel GJ (1998) Modulation of cytokine-induced HIV gene expression by competitive binding of transcription factors to the coactivator p300. *EMBO J* **17**: 3124–3134
- John S, Vinkemeier U, Soldaini E, Darnell Jr JE, Leonard WJ (1999) The significance of tetramerization in promoter recruitment by Stat5. *Mol Cell Biol* **19**: 1910–1918
- Johnson BA, Blevins RA (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**: 604–613
- Kasper LH, Fukuyama T, Biesen MA, Boussouar F, Tong C, de Pauw A, Murray PJ, van Deursen JMA, Brindle PK (2006) Conditional knockout mice reveal distinct functions for the global transcriptional coactivators CBP and p300 in T-cell development. *Mol Cell Biol* **26**: 789–809
- Koradi R, Billeter M, Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics* **14**: 51–55
- Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**: 477–486
- Lee C-H, Saksela K, Mirza UA, Chait BT, Kuriyan J (1996) Crystal structure of the conserved core of the HIV-1 Nef complexed with a Src family SH3 domain. *Cell* **85**: 931–942
- Legge GB, Martinez-Yamout MA, Hambly DM, Trinh T, Lee BM, Dyson HJ, Wright PE (2004) ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. *J Mol Biol* **343**: 1081–1093
- Lodige I, Marg A, Wiesner B, Malecova B, Oelgeschlager T, Vinkemeier U (2005) Nuclear export determines the cytokine sensitivity of STAT transcription factors. *J Biol Chem* **280**: 43087–43099
- Naar AM, Lemon BD, Tjian R (2001) Transcriptional coactivator complexes. *Annu Rev Biochem* **70**: 475–501
- O'Sullivan A, Chang HC, Yu Q, Kaplan MH (2004) STAT4 is required for interleukin-12-induced chromatin remodeling of the CD25 locus. *J Biol Chem* **279**: 7339–7345
- O'Sullivan LA, Liongue C, Lewis RS, Stephenson SE, Ward AC (2007) Cytokine receptor signaling through the Jak-Stat-Socs pathway in disease. *Mol Immunol* **44**: 2497–2506
- Parker D, Ferreri K, Nakajima T, LaMorte VJ, Evans R, Koerber SC, Hoeger C, Montminy M (1996) Phosphorylation of CREB at Ser133 induces complex formation with CPB via a direct mechanism. *Mol Cell Biol* **16**: 694–703
- Pessier F, Cron RQ (2004) Reciprocal regulation of the nuclear factor of activated T cells and HIV-1. *Genes Immun* **5**: 158–167
- Pfützner E, Jahne R, Wissler M, Stoecklin E, Groner B (1998) p300/CREB-binding protein enhances the prolactin-mediated transcriptional induction through direct interaction with the transactivation domain of Stat5, but does not participate in the Stat5-mediated suppression of the glucocorticoid response. *Mol Endocrinol* **12**: 1582–1593
- Ponting CP, Blake DJ, Davies KE, Kendrick-Jones J, Winder SJ (1996) ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem Sci* **21**: 11–13
- Sadzak I, Schiff M, Gattermeier I, Glinitzer R, Sauer I, Saalmüller A, Yang E, Schaljo B, Kovarik P (2008) Recruitment of Stat1 to chromatin is required for interferon-induced serine phosphorylation of Stat1 transactivation domain. *Proc Natl Acad Sci USA* **105**: 8944–8949
- Schuringa JJ, Schepers H, Vellenga E, Kruijer W (2001) Ser727-dependent transcriptional activation by association of p300 with STAT3 upon IL-6 stimulation. *FEBS Lett* **495**: 71–76
- Shen Y, Darnell Jr JE (2001) Antiviral response in cells containing Stat1 with heterologous transactivation domains. *J Virol* **75**: 2627–2633
- Sugase K, Landes MA, Wright PE, Martinez-Yamout M (2008) Overexpression of post-translationally modified peptides in *Escherichia coli* by co-expression with modifying enzymes. *Protein Expr Purif* **57**: 108–115
- Sun W, Snyder M, Levy DE, Zhang JJ (2006) Regulation of Stat3 transcriptional activity by the conserved LPMS motif for OSM and IL-6 signaling. *FEBS Lett* **580**: 5880–5884
- Sun W, Xu W, Snyder M, He W, Ho H, Ivashkiv LB, Zhang JJ (2005) The conserved Leu-724 residue is required for both serine phosphorylation and co-activator recruitment for Stat1-mediated transcription activation in response to interferon- γ . *J Biol Chem* **280**: 41844–41851
- Teufel DP, Freund SM, Bycroft M, Fersht AR (2007) Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proc Natl Acad Sci USA* **104**: 7009–7014
- Tsui V, Case DA (2000) Molecular simulations of nucleic acids using a generalized Born solvation model. *J Am Chem Soc* **122**: 2489–2498
- Vanhatupa S, Ungureanu D, Paakkunainen M, Silvennoinen O (2008) MAPK-induced Ser727 phosphorylation promotes SUMOylation of STAT1. *Biochem J* **409**: 179–185
- Varinou L, Ramsauer K, Karaghiosoff M, Kolbe T, Pfeffer K, Müller M, Decker T (2003) Phosphorylation of the Stat1 transactivation domain is required for full-fledged IFN- γ -dependent innate immunity. *Immunity* **19**: 793–802
- Vinkemeier U, Cohen SL, Moarefi I, Chait BT, Kuriyan J, Darnell Jr JE (1996) DNA binding of *in vitro* activated Stat1 α , Stat1 β and truncated Stat1: interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites. *EMBO J* **15**: 5616–5626
- Vuister GW, Bax A (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J($H^N H^{\alpha}$) coupling constants in ^{15}N -enriched proteins. *J Am Chem Soc* **115**: 7772–7777
- Wen Z, Zhong Z, Darnell Jr JE (1995) Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* **82**: 241–250
- Wittekind M, Mueller L (1993) HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α - and β -carbon resonances in proteins. *J Magn Reson* **101**: 201–205
- Xu X, Sun YL, Hoey T (1996) Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. *Science* **273**: 794–797
- Zakharova N, Lyman ES, Yang E, Malik S, Zhang JJ, Roeder RG, Darnell Jr JE (2003) Distinct transcriptional activation functions of STAT1 $\{\alpha\}$ and STAT1 $\{\beta\}$ on DNA and chromatin templates. *J Biol Chem* **278**: 43067–43073
- Zhang JJ, Vinkemeier U, Gu W, Chakravarti D, Horvath CM, Darnell Jr JE (1996) Two contact regions between Stat1 and CBP/p300 an interferon γ signaling. *Proc Natl Acad Sci USA* **93**: 15092–15096