Structural insights into TAZ2 domain-mediated CBP/p300 recruitment by transactivation domain 1 of the lymphopoietic transcription factor E2A

Marina R. Lochhead[1], Alexandra D. Brown[2], Alyssa C. Kirlin[1], Seth Chitayat[1], Kim Munro[3], Jane E. Findlay[4], George S. Baillie[4], David P. LeBrun[5], David N. Langelaan[1,2,#]*, and Steven P. Smith[1,3]*

From the [1]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON, K7L 3N6, Canada; [2]Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, NS, B3H 4R2, Canada; [3]Protein Function Discovery Group, Queen's University, Kingston, ON, K7L 3N6, Canada; [4]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, G12 8QQ, Scotland ; [5]Department of Pathology and Molecular Medicine, Queen's University, Kingston, ON, K7L 3N6, Canada

Running title: *Structure of the E2A-AD1:TAZ2 complex*

[#]Present address: Dalhousie University, Halifax, NS, Canada, B3H 4R2

*To whom correspondence should be addressed: Steven P. Smith, Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, K7L 3N6, Canada; steven.smith@queensu.ca; Tel:(1)613-533-3188.
David N. Langelaan, Department of Biochemistry & Molecular Biology, Dalhousie University, Halifax, NS, B3H 4R2, Canada; david.langelaan@dal.ca; Tel:(1)902-494-8928.

Keywords: hematopoiesis, lymphocyte, basic helix-loop-helix transcription factor (bHLH), E1A binding protein p300 (p300), protein structure, protein-protein interaction, nuclear magnetic resonance (NMR), isothermal titration calorimetry (ITC), intrinsically disordered protein, TAZ2 domain

ABSTRACT

The E-protein transcription factors guide immune cell differentiation, with E12 and E47 (hereafter called E2A) being essential for B-cell specification and maturation. E2A and the oncogenic chimera E2A–PBX1 contain three transactivation domains (ADs), with AD1 and AD2 having redundant, independent, and cooperative functions in a cell-dependent manner. AD1 and AD2 both mediate their functions by binding to the KIX domain of the histone acetyltransferase paralogues CREB-binding protein (CBP) and E1A-binding protein P300 (p300). This interaction is necessary for B-cell maturation and oncogenesis by E2A–PBX1 and occurs through conserved ϕ-x-x-ϕ-ϕ motifs (with ϕ denoting a hydrophobic amino acid) in AD1 and AD2. However, disruption of this interaction via mutation of the KIX domain in CBP/p300 does not completely abrogate binding of E2A and E2APBX1. Here, we determined that E2A–AD1 and E2A–AD2 also interact with the TAZ2 domain of CBP/p300. Characterization of the TAZ2:E2AAD1(1–37) complex indicated that E2A–AD1 adopts an α-helical structure and uses its ϕ-x-x-ϕ-ϕ motif to bind TAZ2. While this region overlapped with the KIX recognition region, key KIX-interacting E2A–AD1 residues were exposed, suggesting that E2A–AD1 could simultaneously bind both the KIX and TAZ2 domains. However, we did not detect a ternary complex involving E2A–AD1, KIX, and TAZ2 and found that E2A containing both intact AD1 and AD2 is required to bind to CBP/p300. Our findings highlight the structural plasticity and promiscuity of E2A–AD1 and suggest that E2A binds both the TAZ2 and KIX domains of CBP/p300 through AD1 and AD2.

The innate and adaptive immune systems rely on the development and differentiation of hematopoietic stem cells to mature blood cells, including B and T lymphocytes, natural killer cells, and plasmacytoid dendritic cells (1–3). The

generation of these immune cells, or lymphopoiesis, involves numerous intermediates, a progressive limitation of differentiation potential, and a coincident loss of the ability to self-renew (4). This process of lineage commitment is tightly regulated at the transcriptional level by complex regulatory networks comprising both lineage-specific and ubiquitous transcription factors (2, 5–7). One such set of transcription factors is the E-protein family, which comprise class I basic helix-loop-helix (bHLH) transcription factors that play essential roles in the development and specification of B- and T-lymphocytes (8–10).

Members of the E-protein family include the alternatively spliced isoforms E12 and E47 (also referred to collectively as E2A), HEB, and E2-2. Each family member contains a C-terminal bHLH domain responsible for E-protein dimerization and binding DNA at E-box CANNTG consensus sites in gene enhancer and/or promoter regions (11–17). In addition, the E-proteins possess three conserved activation domains, one of which (AD1) is positioned at the extreme N-terminus while the other two (AD2 and AD3) are more centrally located (Fig. 1A) (14, 15, 18, 19). These activation domains have been shown to display cooperative, or independent transcriptional regulatory functions in a cell-specific manner (19–25). For example, AD1, AD2, and AD3 independently induce transcriptional activation but bind the same site of co-activators in a redundant manner, and when combined AD1 and AD2 cooperate to display greater than additive gene induction (19, 21, 22, 26–28). Deletion of AD1 or AD2 in E2A abolishes B-lymphoid differentiation in a pre-B cell line, which is consistent with the observation that E2A plays a critical regulatory role at the earliest stages of B-lymphoid specification (21). The transcriptional regulatory role of these activation domains resides in their ability to recruit general transcriptional factors and coactivators, such as TFIID and the histone acetyltransferases SAGA, GCN5, PCAF, the paralogs CBP and p300, and corepressors such as ETO (19, 22, 23, 26, 29–31). Whereas AD3 allows the E-proteins to recruit TFIID to the core promoter by binding the TAFH domain of TAF4 (19), both AD1 and AD2 interact with CBP/p300 to enhance the acetyltransferase activity of this cofactor (26–28, 32). A conserved region within AD1 called the 'p300/CBP and ETO target in E proteins' (PCET) motif, is also the target of the transcriptional repressor ETO, and competition between ETO and CBP/p300 for binding to the PCET motif has been proposed to be the mechanism underlying E-protein mediated transcriptional silencing (31).

CBP/p300 are multimodular proteins in which several protein-protein interaction domains allow recruitment to enhancers and promoters via interactions with the activation domains of an array of transcription factors (Fig. 1A) (33–35). AD1 and AD2 of the E-proteins have been reported to bind the same surface of the KIX domain of CBP/p300 via their Φ-x-x-Φ-Φ sequences (where Φ represents a hydrophobic amino acid and x any other amino acid), which comprise the core PCET motif (Fig. 1B). This interaction has been shown to be particularly important for leukemia induction by the oncogenic protein E2A-PBX1 that arises from a t(1;19) chromosomal translocation (20, 27, 28). In addition to the KIX domain, the TAZ1 and TAZ2 domains of CBP/p300 have been shown to bind activation domains of several transcription factors, including B-Myb, C/EPBε, FOXO3a, HIF-1α, p53, p63, p73, STAT1, and STAT2; many of which also contain Φ-x-x-Φ-Φ sequences (Fig. 1B) (36–47). Furthermore, several activation domains have displayed binding promiscuity to the KIX, TAZ1, and TAZ2 domains presenting the opportunity for a multivalent mode of binding with transcription factors comprising multiple activation domains; with p53 representing the archetypical example (33, 38, 48, 49). Despite functional roles for the three activation domains being attributed to the transcriptional activity of E-proteins, their interactions with CBP/p300 have not been fully explored.

Here, we characterize a direct interaction between E2A-AD1 and the TAZ2 domain of CBP/p300, and elucidate the structure of a E2A-AD1(1-37):TAZ2 complex by NMR spectroscopy. The structure shows residues throughout and adjacent to the helical PCET motif of E2A-AD1 interact with the TAZ2 domain in a manner reminiscent of activation subdomain 2 (AD2) of p53. Peptide microarray and mutagenesis revealed the requirement of both hydrophobic and electrostatic interactions in complex formation. These results provide a mechanistic rational for the cooperative manner in which the AD1 and AD2 domains of E2A induce gene expression.

Results

TAZ2 domain of CBP/p300 binds AD1 of E2A

Towards identifying which domains of CBP/p300 interact with E2A, and in particular its three activation domains (i.e., AD1, AD2, AD3), an *in vitro* pulldown experiment was performed involving GB1-E2A fusion constructs that spanned the entire transactivation region of E2A (residues 1-483; Fig. 2 and S1). The KIX domain of CBP/p300 interacted modestly with E2A-AD1, and a significant interaction with E2A(1-483) was observed, which is consistent with previous studies ((26–28); Fig. 2). E2A did not significantly interact with TAZ1, but displayed a more pronounced ability to pulldown the isolated TAZ2 domain, as indicated by the corresponding protein bands from pulldowns with E2A-AD1, E2A-AD2 and E2A(1-483) (Fig. 2). A quantitative assessment of the E2A:TAZ2 interaction by isothermal titration calorimetry indicated that the TAZ2 domain displayed a higher affinity for E2A-AD1 (i.e., dissociation constant (K_d) of 0.86 ± 0.04 μM) than for E2A-AD2 (K_d of 67 ± 7 μM), and that the former occurred with a 1:1 stoichiometric ratio (Fig. S2).

Analysis of the E2A-AD1 interaction with TAZ2 by NMR spectroscopy revealed poor dispersion of the backbone amide resonances of uniformly ^{15}N-labeled E2A-AD1, which suggests that E2A is intrinsically disordered (Fig. 3A). Secondary structure propensity (50) analysis indicated that while E2A-AD1 is largely disordered, residues Lys14-Met25 have elevated SSP scores between 0.1 and 0.35, suggesting that this region has residual α-helical structure (Fig. S3). Addition of saturating amounts of unlabeled TAZ2 domain caused chemical shift changes and increased spectral dispersion, in particular for the backbone amide groups of residues Lys14-Phe26 and Leu94 of E2A-AD1 (Fig. 3A insert and S4). Isothermal titration calorimetry measured the dissociation constants of TAZ2 binding for E2A-AD1 and E2A-AD1(1-37) to be 0.86 ± 0.04 μM and 4.1 ± 1.2 μM, respectively (Fig. S2). This is consistent with the chemical shift perturbation data of E2A-AD1, which indicates that the key TAZ2 interacting region of E2A-AD1 lies within E2A-AD1(1-37), while minor contributions from elsewhere in E2A-AD1 may lead to increased TAZ2 binding affinity for E2A-AD1 compared to E2A-AD1(1-37).

The interaction of E2A-AD1(1-37) with TAZ2 was further characterized by NMR spectroscopy because it is more feasible to characterize smaller proteins. Similar to what was observed for E2A-AD1, addition of saturating amounts of unlabeled TAZ2 domain to ^{15}N labeled E2A-AD1(1-37) caused large chemical shift changes and increased spectral dispersion for resonances corresponding to the backbone amide groups of Lys14-Phe26 (Fig. 3B; red vs. black spectra). Chemical shift analysis of backbone ^{1}H, ^{13}C, and ^{15}N chemical shifts of free and TAZ2-bound E2A-AD1(1-37) suggest that Asp13-Met25 undergo a random coil to α-helix structural change upon TAZ2 binding (Fig. S5). Steady-state $\{^{1}H\}$-^{15}N heteronuclear NOE analysis further indicated that when bound to the TAZ2 domain, E2A-AD1(1-37) adopted a more ordered structure between Lys14 and Leu28 (i.e., $\{^{1}H\}$-^{15}N NOE values > 0.4) while the preceding N-terminal residues and the seven C-terminal residues displayed reduced $\{^{1}H\}$-^{15}N NOE values suggestive of a disordered conformation in solution (Fig. 3C).

Structure of the E2A-AD1(1-37):TAZ2 complex

To investigate the molecular basis of the E2A-AD1(1-37):TAZ2 interaction, the structure of this complex was determined by NMR spectroscopy using 3408 NOE-derived distance restraints, including 197 intermolecular distance restraints, and 163 TALOS+ derived dihedral restraints. Structures were calculated from automated distance restraint weighting and assignment through CYANA and water refinement with CNS (51–53). The final ensemble of 20 low-energy structures of the 1:1 complex was well defined (Fig. 4A) and displayed good overall structural statistics (Table 1), with root mean square deviation (r.m.s.d.) values of the ensemble core (E2A-AD1(1-37) residues Asp13-Pro29 and TAZ2 residues Ser1726-Lys1812) from the energy-minimized average structure for the backbone and heavy atoms of 0.51 ± 0.06 Å and 0.83 ± 0.06 Å, respectively.

The E2A-AD1(1-37) adopted an α-helix comprising amino acid residues Gly11-Phe26 while Pro27-Pro29 formed a short extended region (Fig. 4). The ten N-terminal amino residues (Met1-Val10) and eight C-terminal residues (Val30-Pro37) remained disordered, flexible, and did not appreciably contact the TAZ2 domain, as indicated by the reduced $\{^{1}H\}$-^{15}N NOE values (Fig. 3C) and lack of identifiable medium- and long-range NOEs.

When compared to the structure of HEB-AD1 bound to the KIX domain (27), E2A-AD1(1-37) was extended by one helical turn at the N-terminus (residues Gly11-Lys14).

The TAZ2 domain comprised four α-helices (α1: Pro1727-Gln1747; α2: Pro1756-Lys1769; α3: Pro1780-His1795; α4: Pro1804-Lys1810) separated by loops (Fig. 4), each of which adopted a HCCC type Zn^{2+}-coordinating conformation (54). Overall, the TAZ2 fold in the presence of E2A-AD1(1-37) described here was similar to that previously observed for this domain when bound to other proteins, including the E1A oncoprotein (backbone r.m.s.d of 1.9 Å), isolated p53-AD1 and -AD2 (backbone r.m.s.d. of 2.4 Å and 2.0 Å, respectively), the N-terminal region of p53 comprising both ADs (backbone r.m.s.d. of 1.8 Å), STAT1-AD (backbone r.m.s.d. of 1.5 Å), p63 (backbone r.m.s.d. of 2.0 Å) and p73 (backbone r.m.s.d. of 1.9 Å) (38, 39, 41, 42, 44, 47).

E2A-AD1:TAZ2 interface

The E2A-AD1 interactive surface on the TAZ2 domain involved a hydrophobic groove bounded by basic residues and covered ~1260 $Å^2$ of solvent accessible surface area (Fig. 5A,B). The hydrophobic groove comprised the aliphatic region of Arg1732, Ile1735 and Ala1738 of α1, the aliphatic region of Lys1760, Met1761 and Val1764 of α2, and Pro1780, Ile1781, Gln1784, Leu1785, Leu1788 and Tyr1791 of α3 (Fig. 5C). Residues throughout the helical region of E2A-AD1, including Leu19, Phe22, Met25 and Phe26, and the residues Pro27 and Leu28 C-terminal to the helix made extensive non-polar contacts with the TAZ2 domain (Fig. 5D). Within the E2A-AD1 helix, the side chain of Leu19 inserted into a cavity formed by Val1764, Pro1780, Ile1781, the aliphatic region of Gln1784, and Leu1785. The aromatic side chain of Phe22 from E2A-AD1 participated in van der Waals contacts with the aliphatic portion of Lys1760, Met1761, and Val1764. Met25 interacted with Ala1738 and Leu1788 while Phe26 associated with Ile1735, Ala1738, Ala1787, Leu1788, and the aliphatic region of Gln1784. Beyond the helical region of E2A-AD1, Pro27 formed non-polar contacts with Ile1735 while Leu28 resided in hydrophobic pocket created by aliphatic region of Arg1732, Ile1735, Leu1788, and Tyr1791 of the TAZ2 domain.

Several electrostatic and hydrogen bonding interactions involving residues in the helical region of E2A-AD1 complemented the extensive non-polar network at the E2A-AD1:TAZ2 interface. Glu15 of E2A-AD1 was situated to form a salt bridge with His1767 of α2 of TAZ2 ($d_{Oε*\text{-}Nδ1}$ = 3.6 ± 0.4 Å). Asp18 was positioned to participated in salt bridges with both Lys1760 ($d_{Oδ*\text{-}Nζ}$ = 3.1 ± 0.5 Å) and Arg1763 ($d_{Oδ2\text{-}Nε}$ = 4.5 ± 1.5 Å) while Asp21 was also in close proximity to Lys1760 of TAZ2 ($d_{Oδ*\text{-}Nζ}$ = 3.6 ± 0.9 Å).

Mutants traversing the helical region of E2A-AD1 attenuate TAZ2 binding

To assess the relative contribution of each residue of E2A-AD1 in TAZ2 recognition, the effect of E2A-AD1 alanine mutants were assessed by peptide microarray and isothermal titration calorimetry (Fig. 6). For the peptide microarray, HEB-AD1(11-28) peptides (which is identical to E2A-AD1 except for a Met24 to Ala substitution), in which each position was individually substituted to alanine, were probed with recombinant His_6-GB1-TAZ2 (Fig. 6A). Compared to wild-type HEB-AD1(11-28), alanine substitutions at Leu19 and Phe22 led to significantly reduced signal while substitutions at Leu16, Asp18, and Asp21 modestly reduced the signal. Complementary quantitative analysis by isothermal titration calorimetry indicated that the Leu19Ala and Phe22Ala substitutions both resulted in a ~4-fold reduction in affinity (K_d of 17 μM and 16 μM, respectively) (Fig. 6B). Thus, while no single point mutation of E2A-AD1 completely abrogates TAZ2 binding, the E2A-AD1(1-37):TAZ2 structure, peptide microarray, and isothermal titration calorimetry indicate that L19 and F22 are key binding elements of the interaction.

KIX and TAZ2 of CBP/p300 compete for E2A-AD1

The orientation of E2A-AD1 on the TAZ2 surface presented the possibility of a higher-order interaction with the KIX domain of CBP/p300 as several of the KIX-interactive residues, including the critical Leu20 (20), were solvent exposed; a scenario supported by the recent observation that full-length CBP/p300 displays intrinsic conformational flexibility (33, 55). To directly assess the formation of such a complex or whether KIX and TAZ2 displayed exclusive E2A-AD1

binding, an NMR-based displacement experiment was performed. Addition of unlabeled E2A-AD1(1-37) to uniformly ^{15}N-labeled KIX resulted in significant chemical shift changes for a subset of backbone amide resonances in ^{15}N-KIX spectra (Fig. 7A), consistent with the site-specific binding of E2A-AD1(1-37) to KIX (27). Subsequent addition of unlabeled TAZ2 to the sample caused those backbone resonances of uniformly ^{15}N-labeled KIX to revert back towards chemical shift values consistent with the free form of the KIX domain (Fig. 7B). These results indicate that the TAZ2 and KIX domains of CBP/p300 can only interact with E2A-AD1 exclusive of one another.

To further investigate a higher-order interaction between CBP and E2A, a pulldown experiment involving GB1-E2A(1-483) constructs and full-length FLAG-tagged CBP was performed and visualized using an anti-FLAG western blot (Fig. 8 and S6). Compared to wild-type, alanine substitutions in the AD1 region of E2A(1-483) that disrupt either TAZ2 binding, (Leu19Ala and Phe22Ala) or KIX binding (Leu20Ala; (20)) significantly reduced the ability of E2A(1-483) to pulldown CBP. An alanine and proline substitution in AD2 of E2A(1-483) that disrupts KIX binding (Leu397Ala/Ile401Pro; (28)) similarly displayed reduced pulldown of flag-CBP. Mutations in both AD1 and AD2 of E2A(1-483) were required to abrogate any observable interaction between GB1-E2A(1-483) and CBP.

Discussion

The E2A activation domains are essential for lymphopoiesis and are involved in the onset of ALL through the oncogenic fusion protein E2A-PBX1 (56). In either situation, the recruitment of CBP/p300 is essential for E2A or E2A-PBX1 function. Here we investigate recruitment of CBP/p300 by E2A-AD1, which provides insight into the molecular mechanisms underlying lymphopoiesis and ALL.

Depending on the context, the activation domains of E2A function independently, redundantly, or cooperatively with each other (19, 21–28, 57). The KIX domain of CBP/p300 binds both E2A-AD1 and E2A-AD2 in a functionally redundant manner, with both E2A-AD1 and E2A-AD2 competing for the same site of KIX (27, 28). The additional interactions of TAZ2 with E2A-AD1 or E2A-AD2 that we describe (Fig. 2 and S2) raise the possibility that both AD1 and AD2 could simultaneously interact with the KIX and TAZ2 domains to allow for higher affinity association between CBP/p300 and E2A. This possibility is supported by our pulldowns of full-length CBP (Fig. 8), where intact AD1 and AD2 were required for maximum CBP/p300 pulldown and mutations in either the AD1 or AD2 region lessened the interaction by ~50%. It is unclear what the binding preferences of E2A-AD1 and E2A-AD2 are to full length CBP since E2A-AD1 and E2A-AD2 are able to bind both the KIX and TAZ2 domains. The ability of E2A to form a tight yet dynamic complex with CBP through multiple weaker interactions is likely essential for its transcriptional activity and provides rationale for how deletion of AD1 or AD2 abolishes B-lymphoid differentiation (21). This is reminiscent of p53, which also associates with multiple domains of CBP/p300 to form a tight, yet dynamic, complex (49).

The promiscuous interactions of intrinsically disordered proteins are recognized as key to their roles as protein interaction hubs (33, 58). This appears to be the case with E2A since E2A-AD1 and its nearly identical homologue HEB-AD1 (Fig. 1B) can complement and bind a variety of molecular surfaces, with high-resolution structures available of the KIX:HEB-AD1 and eTAFH:HEB-AD1 complexes. HEB-AD1 adopts an amphipathic α-helix from Lys14-Phe26 (renumbered according to the E2A sequence) when bound to KIX, a kinked α-helix spanning residues Asp13-Leu20 and Ser23-Phe26 when bound to eTAFH, and E2A-AD1 adopts an amphipathic α-helix from Asp13-Met25 when bound to TAZ2. Leu16, Leu19, and Leu20 form essential hydrophobic contacts with both eTAFH and KIX (27, 59), while Phe22 forms additional hydrophobic contacts with KIX and TAZ2 but not eTAFH. Although the entire PCET motif is involved in recognition of eTAFH, KIX, and TAZ2, Phe26, Pro27 and Leu28 have additional hydrophobic contacts to TAZ2. Finally eTAFH, KIX, and TAZ2 have basic residues situated nearby the E2A-AD1 binding site, providing unique potential electrostatic contacts for Glu15, Asp18 and Asp21. The different modes of E2A-AD1 binding to eTAFH, KIX, and TAZ2 are highlighted by the importance of Leu20, which is essential for the HEB:KIX, and HEB:eTAFH interactions (27, 59), but dispensable for binding to TAZ2 as it is facing solvent and alanine

substitutions are non-perturbing to a pulldown experiment (Fig. 5 and 6).

The TAZ2 domain of CBP/p300 is promiscuous and interacts with many intrinsically disordered proteins, with structures available for TAZ2 in complex with activation domains from E1A, C/EBPε, p53, p63, p73, and STAT 1 (36, 38, 39, 41, 42, 47). Highlighting the promiscuity of the TAZ2 domain, E1A, p73, and STAT 1 bind TAZ2 in part to a hydrophobic groove formed by α1, α2, and α3, while p53, p63, and E2A-AD1 adopt an α-helical structure when they bind to this same region of TAZ2 (Fig. 9). The adenoviral protein E1A has been shown to decrease E2A transcriptional activity (60) and the observation that E2A-AD1 and E1A bind overlapping sites of TAZ2 suggests that E1A inhibits E2A function through direct competition for TAZ2. Interestingly, although E2A-AD1, p53, and p63 bind the same region of TAZ2 as an α-helix, the peptides have differing orientations (Fig. 9 and 10). Despite the opposite orientations of p53 and E2A-AD1, Phe22 and Phe26 of E2A-AD1 occupy similar positions as Phe54 and Ile50, respectively (Fig. 10). These residues all face TAZ2, suggesting that an amphipathic helix is key for binding to the α1, α2 and α3 binding surface of TAZ2. The importance of these hydrophobic residues is confirmed by our mutagenesis studies in which mutation of Phe22 or Phe26 to alanine decreased the signal observed in a microarray, while mutation of I50 and Phe54 of p53 weakens binding of p53-AD2 to TAZ2 by approximately 3 fold and 2 fold, respectively (44). Other than the presence of a Φ-x-x-Φ-Φ sequence and being acidic with some hydrophobic residues, the sequences of TAZ2-binding activation domains are quite divergent and TAZ2 appears to be able to accommodate a variety of peptide sequences, orientations, and secondary structure content.

The orientation of an amphipathic α-helix bound to TAZ2 likely depends on the specific interactions of polar residues and neighbouring sequences with TAZ2. For E2A-AD1 Asp18 and Asp21 are likely involved in a salt bridge to the sidechain of Lys1760 ($N^{\zeta} - O^{\delta}$ distance of 3.1 ± 0.53 Å and 3.6 ± 0.9 Å, respectively), while Phe26, Pro27 and Leu28 all participate in further hydrophobic contacts to TAZ2. For p53 a similar situation occurs where Glu11 and Glu17 contact the guanidinium group of Arg1731 and multiple other residues engage in specific polar contacts to TAZ2 (41). In support of this, substitution of Asp18 and Asp21 with alanine decreases TAZ2 binding in a peptide microarray (Fig. 6). Given that acidic residues influence E2A-AD1 affinity for TAZ2, phosphorylation of E1A (e.g. at Thr12, Ser17 and Ser23) may enhance the affinity of E2A-AD1 to TAZ2 by forming salt bridges to nearby basic residues of TAZ2. A similar situation has been observed with p53, which has a graded enhancement of affinity for CBP/p300 upon phosphorylation (61). Overall, TAZ2 is able to accommodate a wide variety of peptides through a complex interplay of hydrophobic and electrostatic interactions, and phosphorylation may be a common method to regulate protein association with TAZ2.

Experimental procedures

Plasmid preparation. All constructs were cloned into a pET21a(+) vector using *Bam*HI and *Xho*I restriction enzymes downstream of sequences coding for a hexahistidine (His_6) tag, the B1 domain of protein G (GB1) and either a thrombin (human E2A(1-37) (hereafter referred to as E2A-AD1(1-37)), CBP(346-440) (TAZ1), and p300(1723-1812) (TAZ2)) or TEV (human E2A(1-100) (E2A-AD1), E2A(101-300) (E2A-AD3), E2A(301-483) (E2A-AD2), and E2A(1-483)) protease recognition sequence. A modified TAZ2 construct in which cysteine residues not involved in zinc coordination were substituted for alanine (i.e., Cys1738Ala, Cys1746Ala, Cys1789Ala, Cys1790Ala) was generated based on its reported enhanced stability (41, 54). E2A-AD1(1-37) mutants were generated using the QuikChange site-directed mutagenesis kit (Stratagene), with His_6-GB1-E2A-AD1(1-37) as a template. E2A(1-483) mutagenesis was performed using QuikChange II site-directed mutagenesis kit (Agilent Technologies, 200523) with His_6-GB1-E2A(1-483) as a template. The fidelity of all constructs was verified by DNA sequencing. Plasmids were subsequently transformed into *Escherichia coli* BL21(DE3) cells for recombinant protein expression.

Protein expression and purification. The isolated KIX domain (residues 586-673 of CBP) was expressed and purified as previously described (27). *E. coli* BL21(DE3) cells harbouring the His_6-GB1-TAZ1 and His_6-GB1-TAZ2 encoding plasmids were grown in LB or ^{15}N- or $^{13}C/^{15}N$-enriched M9 media at 37°C supplemented with 100

μM $ZnCl_2$. Protein expression was induced at optical density of 0.6 at 600 nm by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Growth was continued overnight at 23°C with shaking. Harvested cell pellets were lysed by sonication in denaturing binding buffer (20 mM Tris-HCl pH 8.0, 250 mM NaCl, 8 M urea, 10 mM β-mercaptoethanol, 10 μM $ZnCl_2$), clarified by centrifugation, and applied to Ni^{2+}-affinity resin (GE Healthcare). Upon extensive washes with denaturing binding buffer containing 10 mM imidazole, protein constructs were refolded on-column by application of native Ni^{2+} binding buffer (20 mM Tris-HCl pH 8.0, 250 mM NaCl, 10 mM β-mercaptoethanol, 10 μM $ZnCl_2$), and subsequently eluted in the same buffer containing 300 mM imidazole. The elution fractions were pooled, and β-mercaptoethanol and $ZnCl_2$ were added to final concentrations of 40 mM and four-fold excess relative protein, respectively. Thrombin, at 1 unit per 50 nanomoles of protein, was added and the samples were dialyzed overnight in buffer A (20 mM HEPES pH 8.0, 50 mM NaCl, 5 mM β-mercaptoethanol, 10 μM $ZnCl_2$). The cleaved, refolded TAZ1 or TAZ2 constructs were separated from the His_6-GB1 fragments via fast flow SP sepharose cation chromatography (GE Healthcare) using buffer A as a wash buffer, and eluted with buffer A containing 500 mM NaCl.

Expression of all His_6-GB1-E2A constructs were induced in transformed *E. coli* BL21(DE3) cells with 0.5 mM IPTG at an optical density of ~1 at 600 nm. Growth was continued for an additional 4 hours at 37°C with shaking. Expression of His_6-GB1-E2A(1-483) L397A/I401P mutants was performed as described except that after induction with IPTG, growth was continued for an additional 16 hours at 23°C with shaking. Purification of His_6-GB1-E2A(1-483), E2A-AD1, and E2A-AD2 was performed as described above for the TAZ domains, with the exception that $ZnCl_2$ was excluded from all buffers. Purification of wild-type and mutant E2A-AD1(1-37) constructs was performed as previously described (28).

For uniformly ^{13}C- and/or ^{15}N-labeled NMR samples high performance liquid chromatography was used to purify TAZ2 or E2A-AD1(1-37) on a C_{18} reverse phase column with a water:acetonitrile gradient with 0.05% trifluoroacetic acid. Fractions containing protein were pooled, lyophilized, and stored at -20°C. The integrity of each protein sample was verified by SDS-PAGE analysis, mass spectrometry, and NMR spectroscopy.

***In vitro* pulldown experiments**. His_6-GB1, His_6-GB1-E2A-AD1, His_6-GB1-E2A-AD3, His_6-GB1-E2A-AD2, and His_6-GB1-E2A(1-483) were incubated with 20 μL of 50% IgG agarose slurry (GE Healthcare) for 15 minutes in assay buffer (20 mM MES pH 6.5, 100 mM NaCl). After 2 washes with assay buffer, 1 mL of 20 μM KIX, TAZ1 or TAZ2 was added to the beads and left for 30 minutes with gentle agitation. After incubation, the beads were washed three times with assay buffer and resuspended in SDS-PAGE loading buffer. KIX, TAZ1, and TAZ2 binding to the His_6-GB1-E2A constructs were visualized using SDS-PAGE. For TAZ1 pulldown experiments, the assay buffer included 5 mM β-mercaptoethanol and for TAZ2 pulldown experiments, the assay buffer included 5 mM β-mercaptoethanol and 10 μM $ZnCl_2$. All experiments were done in duplicate.

Cell culture and cell lysis. HEK 293T cells were seeded at 0.8 x 10^6 cells per well in a 6-well culture plate. The following day, the cells were transfected with 2 ug of a flag-CBP plasmid using jetPRIME transfection reagent (Polyplus, 114-01) according to the manufacturers protocol. After 24 hours, the transfection medium was exchanged for cell growth medium. Following 48 hours after transfection, the cells were washed twice with PBS and lysed for 10 minutes at room temperature with gentle agitation in NP-40 buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1.0% NP-40) plus protease inhibitors (Thermo Fisher Scientific, A32965). The lysates were sonicated briefly and clarified by centrifugation.

Immunoblotting. *In vitro* pulldown experiments were performed as described above with the following modifications: i) His_6-GB1 and His_6-GB1-E2A(1-483) wild-type and mutant constructs were incubated with 5 uL of 50% IgG agarose slurry (GE Healthcare) for 30 minutes in NP-40 buffer, ii) 250 uL of flag-CBP transfected HEK 293T cell lysate was added to the beads and left for 3 hours with gentle agitation, iii) all washes were done with NP-40 buffer. The pull-down experiments were loaded onto 6% Tris-glycine gels and separated by SDS-PAGE. The proteins were transferred to nitrocellulose membranes using Bio-

Rad wet electroblotting system for 16 hours at 20 V at 4°C. Membranes were blocked for 1 hour at room temperature with 5% skimmed milk in Tris-buffered saline plus 0.1% Tween 20 (TBST) and subsequently incubated for 1 hour at room temperature with horse radish peroxidase-conjugated α-FLAG antibody (Sigma Aldrich, A8592) diluted 1:5000 in TBST. Membranes were then washed four times with TBST, stained with Immobilin Forte Western HRP substrate (EMD Millipore, WBLUF0100), and exposed to X-ray films. The pull-down experiments and immunoblotting were performed in triplicate and ImageJ was used to quantify the X-ray films.

Isothermal titration calorimetry. Experiments were carried out at 30°C in 20 mM MES pH 6.5, 5 mM β-mercaptoethanol and 10 μM $ZnCl_2$, using 100-200 μM TAZ2 in the cell and 1-2 mM E2A construct in the syringe (E2A-AD1, E2A-AD2, E2A-AD1(1-37), E2A-AD1(1-37)L19A, or E1A-AD1(1-37)F22A). Experiments were performed on VP-ITC or iTC200 calorimeters (MicroCal). Thermograms were fit to a one-site binding model using MicroCal Origin 7.0 software. All experiments were collected in at least duplicate.

Peptide microarray. Peptide microarrays were synthesized through automatic SPOT synthesis with a MultiPep automated peptide synthesis system (Intavis) using 9-fluorenylmethyloxycarbonyl chemistry (62). HEB-AD1 derived peptides were synthesized onto continuous cellulose membranes to generate strip arrays, which were hydrated with ethanol, washed with assay buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl) and blocked overnight in assay buffer containing 2% milk. Following washes with assay buffer containing 0.5% Tween 20, the arrays were incubated for 4 hours with 5-10 μg/mL His_6-GB1-TAZ2 in 50 mM Tris-HCl pH 7.4, in assay buffer containing 2% skim milk, washed with assay buffer, and incubated for 1 hour with rabbit polyclonal antibody against the His_6 tag conjugated to horse radish peroxidase (Abcam) diluted in assay buffer containing 2% skim milk. After a final wash with assay buffer, luminata horseradish peroxidase substrate (Millipore) was applied to the arrays and X-ray film was used to detect the signal.

NMR spectroscopy and structure calculation. Lyophilized recombinant TAZ2 was reconstituted for NMR studies as previously described (54). For the assignment of E2A-AD1, a 1 mM sample of $^{13}C/^{15}N$ labeled E2A-AD1 in 20 mM MES, 50 mM NaCl, and 5% D_2O adjusted to pH 6.5 was prepared, and data was collected at 25°C. For structural analysis of the E2A-AD1(1-37):TAZ2 complex, two samples were prepared containing either 1 mM uniformly $^{13}C/^{15}N$-labeled TAZ2 with 3 mM E2A-AD1(1-37) or 2 mM TAZ2 with 1.4 mM uniformly $^{13}C/^{15}N$-labeled E2A-AD1(1-37) in 20 mM MES, 5 mM β-mercaptoethanol, 100 μM $ZnCl_2$ and 5% D_2O adjusted to pH 6.5 without accounting for deuterium isotope effects. Resonance assignments of E2A-AD1, E2A-AD1(1-37), and TAZ2 were carried out using standard multi-dimensional heteronuclear NMR experiments. Distance restraints for the E2A-AD1(1-37):TAZ2 complex were determined using 3D ^{15}N-NOESY-HSQC and both aliphatic and aromatic ^{13}C-NOESY-HSQC experiments, all with a 100 ms mixing time. The validity of intermolecular restraints was confirmed using $^{12}C/^{14}N$-filtered ^{15}N-edited and $^{12}C/^{14}N$-filtered ^{13}C-edited NOESY spectra. All NMR experiments involving E2A-AD1(1-37) were collected at 15°C on 500, 600 or 800 MHz Varian INOVA spectrometers equipped with triple resonance cryoprobes. NMR data were processed using NMRPipe (63) and analyzed using CCPNMR Analysis (64).

NOESY peaks lists were exported from CCPNMR Analysis and used in CYANA for automatic NOE assignment and distance restraint calibration (52, 53). Dihedral angle restraints based on chemical shifts were generated using TALOS+, with angle error set to two times the error output of TALOS+ (65). After confirming a correct TAZ2 fold using only NOE-based distance and dihedral angle restraints, hydrogen bond restraints ($1.8 \leq d_{OH} \leq 2.2$ Å; $2.7 \leq d_{ON} \leq 3.2$ Å) were applied to helical regions of the E2A-AD1(1-37):TAZ2 complex while zinc coordination restraints were applied to known Zn^{2+} coordinating residues (54). Using CYANA, an initial ensemble of the lowest energy 20 models was retained from 100 generated models. These 20 models were further energy minimized in explicit water using fmcGUI and CNS (51). The final ensemble of 20 models was validated using PROCHECK-NMR (66) and the recall, precision,

F-measure and discriminating power scores of the final ensemble of the E2A-AD1(1-37):TAZ2 complex were calculated using CCPNMR Analysis (64, 67). The protein structure validation software suite was used to determine the ordered residues and evaluate the quality of the final E2A-AD1(1-37):TAZ2 ensemble (68). The ensemble of 20 lowest-energy structural models was deposited to the Protein Data Bank (accession no. 2MH0) while ^{1}H, ^{13}C, and ^{15}N chemical shifts and restraints were deposited to the BMRB (accession no: 19610).

For the NMR-based experiment assessing the ability of TAZ2 and KIX to compete for binding to E2A-AD1 a sample of 200 μM ^{15}N-labeled KIX and 300 μM unlabeled E2A-AD1(1-37) in 20 mM MES pH 6.8, 5 mM β-mercaptoethanol, 10 μM $ZnCl_2$, 95% H_2O/5% D_2O was used. Unlabeled TAZ2 was subsequently added to the sample to a final concentration of 500 μM.

Accession codes. Chemical shift assignments for the E2A-AD1(1-37):TAZ2 complex and E2A-AD1 have been deposited in the BioMagResBank (accession numbers 19610 and 50196, respectively). The atomic coordinates (accession code 2MH0) have been deposited in the Protein Data Bank.

Acknowledgements: We thank Mr. Kim Munro (Protein Function Discovery Facility, Queen's University) for technical assistance in collecting and analyzing the calorimetric data. We thank Dr. Tara Sprules (QANUC NMR Facility, McGill University) and Dr. Ian Burton (NRC, Halifax NS) for assistance with NMR data collection.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

REFERENCES

1. Bryder, D., and Sigvardsson, M. (2010) Shaping up a lineage - lessons from B lymphopoesis. *Curr Opin Immunol*. **22**, 148–153
2. Koch, U., and Radtke, F. (2011) Mechanisms of T cell development and transformation. *Annu. Rev. Cell Dev. Biol.* **27**, 539–562
3. Ramirez, K., and Kee, B. L. (2010) Multiple hats for natural killers. *Curr Opin Immunol*. **22**, 193–198
4. Ikawa, T. (2014) Genetic and epigenetic control of early lymphocyte development. *Curr Top Microbiol Immunol*. **381**, 1–20
5. Hesslein, D. G., and Lanier, L. L. (2011) Transcriptional control of natural killer cell development and function. *Adv. Immunol.* **109**, 45–85
6. Nutt, S. L., and Kee, B. L. (2007) The Transcriptional Regulation of B Cell Lineage Commitment. *Immunity*. **26**, 715–725
7. Santos, P., Arumemi, F., Park, K. S., Borghesi, L., and Milcarek, C. (2011) Transcriptional and epigenetic regulation of B cell development. *Immunol. Res.* **50**, 105–112
8. Braunstein, M., and Anderson, M. K. (2012) HEB in the spotlight: Transcriptional regulation of T-cell specification, commitment, and developmental plasticity. *Clin. Dev. Immunol.* **2012**, 678705
9. de Pooter, R. F., and Kee, B. L. (2010) E proteins and the regulation of early lymphocyte development. *Immunol Rev*. **238**, 93–109
10. Rothenberg, E. V (2014) Transcriptional control of early T and B cell developmental choices. *Annu Rev Immunol*. **32**, 283–321
11. Aronheim, A., Shiran, R., Rosen, A., and Walker, M. D. (1993) Cell-specific expression of helix-loop-helix transcription factors encoded by the E2A gene. *Nucleic Acids Res.* **21**, 1601–1606
12. Greenbaum, S., and Zhuang, Y. (2002) Identification of E2A target genes in B lymphocyte development by using a gene tagging-based chromatin immunoprecipitation system. *Proc Natl Acad Sci U S A*. **99**, 15030–15035
13. Henthorn, P., Kiledjian, M., and Kadesch, T. (1990) Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. *Science*. **247**, 467–470
14. Hu, J. S., Olson, E. N., and Kingston, R. E. (1992) HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. *Mol. Cell. Biol.* **12**, 1031–1042
15. Kee, B. L. (2009) E and ID proteins branch out. *Nat. Rev. Immunol.* **9**, 175–184
16. Massari, M. E., and Murre, C. (2000) Helix-Loop-Helix Proteins: Regulators of Transcription in Eucaryotic Organisms. *Mol. Cell. Biol.* **20**, 429–440
17. Murre, C., McCaw, P. S., Vaessin, H., Caudy, M., Jan, L. Y., Jan, Y. N., Cabrera, C. V, Buskin, J. N., Hauschka, S. D., Lassar, A. B., Weintraub, H., and Baltimore, D. (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*. **58**, 537–544
18. Aronheim, A., Shiran, R., Rosen, A., and Walker, M. D. (1993) The E2A gene product contains two separable and functionally distinct transcription activation domains. *Proc. Natl. Acad. Sci.* **90**, 8063–8067
19. Chen, W.-Y. Y., Zhang, J., Geng, H., Du, Z., Nakadai, T., and Roeder, R. G. (2013) A TAF4 coactivator function for E proteins that involves enhanced TFIID binding. *Genes Dev.* **27**, 1596–1609
20. Bayly, R., Murase, T., Hyndman, B. D., Savage, R., Nurmohamed, S., Munro, K., Casselman, R., Smith, S. P., and LeBrun, D. P. (2006) Critical role for a single leucine residue in leukemia induction by E2A-PBX1. *Mol. Cell. Biol.* **26**, 6442–6452
21. Bhalla, S., Spaulding, C., Brumbaugh, R. L., Zagort, D. E., Massari, M. E., Murre, C., and Kee, B. L. (2008) Differential Roles for the E2A Activation Domains in B Lymphocytes and Macrophages. *J. Immunol.* **180**, 1694–1703
22. Bradney, C., Hjelmeland, M., Komatsu, Y., Yoshida, M., Yao, T.-P. P., and Zhuang, Y. (2003)

Regulation of E2A activities by histone acetyltransferases in B lymphocyte development. *J. Biol. Chem.* **278**, 2370–2376

23. Massari, M. E., Grant, P. A., Pray-Grant, M. G., Berger, S. L., Workman, J. L., and Murre, C. (1999) A Conserved Motif Present in a Class of Helix-Loop-Helix Proteins Activates Transcription by Direct Recruitment of the SAGA Complex. *Mol. Cell*. **4**, 63–73
24. Massari, M. E., Jennings, P. A., and Murre, C. (1996) The AD1 transactivation domain of E2A contains a highly conserved helix which is required for its activity in both Saccharomyces cerevisiae and mammalian cells. *Mol. Cell. Biol.* **16**, 121–129
25. Quong, M. W., Massari, M. E., Zwart, R., and Murre, C. (1993) A new transcriptional-activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells. *Mol. Cell. Biol.* **13**, 792–800
26. Bayly, R., Chuen, L., Currie, R. a, Hyndman, B. D., Casselman, R., Blobel, G. a, and LeBrun, D. P. (2004) E2A-PBX1 interacts directly with the KIX domain of CBP/p300 in the induction of proliferation in primary hematopoietic cells. *J. Biol. Chem.* **279**, 55362–55371
27. Denis, C. M., Chitayat, S., Plevin, M. J., Wang, F., Thompson, P., Liu, S., Spencer, H. L., Ikura, M., LeBrun, D. P., and Smith, S. P. (2012) Structural basis of CBP/p300 recruitment in leukemia induction by E2A-PBX1. *Blood*. **120**, 3968–3977
28. Denis, C. M., Langelaan, D. N., Kirlin, A. C., Chitayat, S., Munro, K., Spencer, H. L., Lebrun, D. P., and Smith, S. P. (2014) Functional redundancy between the transcriptional activation domains of E2A is mediated by binding to the KIX domain of CBP/p300. *Nucleic Acids Res.* **42**, 7370–7382
29. Holmlund, T., Lindberg, M. J., Grander, D., and Wallberg, A. E. (2012) GCN5 acetylates and regulates the stability of the oncoprotein E2A-PBX1 in acute lymphoblastic leukemia. *Leukemia*. **27**, 1–8
30. Scheele, J. S., Kolanczyk, M., Gantert, M., Zemojtel, T., Dorn, A., Sykes, D. P. B., Sykes, D. P. B., Möbest, D. C. C., Kamps, M. P., Räpple, D., Duchniewicz, M., Mobest, D. C., Kamps, M. P., and Rapple, D. (2009) The Spt-Ada-Gcn5-acetyltransferase complex interaction motif of E2a is essential for a subset of transcriptional and oncogenic properties of E2a-Pbx1. *Leuk. Lymphoma*. **50**, 816–828
31. Zhang, J. (2004) E Protein Silencing by the Leukemogenic AML1-ETO Fusion Protein. *Science*. **305**, 1286–1289
32. Hyndman, B. D., Thompson, P., Bayly, R., Côté, G. P., and LeBrun, D. P. (2012) E2A proteins enhance the histone acetyltransferase activity of the transcriptional co-activators CBP and p300. *Biochim. Biophys. Acta*. **1819**, 446–453
33. Dyson, H. J. (2016) Role of intrinsic protein disorder in the function and interactions of the transcriptional protein (CBP) and p300. *J. Biol. Chem.* **291**, 6714–6722
34. Goodman, R. H., and Smolik, S. (2000) CBP/p300 in cell growth, transformation, and development. *Genes Dev*. **14**, 1553–1577
35. Wang, F., Marshall, C. B., and Ikura, M. (2013) Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. *Cell. Mol. Life Sci.* **70**, 3989–4008
36. Bhaumik, P., Davis, J., Tropea, J. E., Cherry, S., Johnson, P. F., and Miller, M. (2014) Structural insights into interactions of C/EBP transcriptional activators with the Taz2 domain of p300. *Acta Crystallogr. D. Biol. Crystallogr.* **70**, 1914–1921
37. Burge, S., Teufel, D. P., Townsley, F. M., Freund, S. M., Bycroft, M., and Fersht, A. R. (2009) Molecular basis of the interactions between the p73 N terminus and p300: effects on transactivation and modulation by phosphorylation. *Proc Natl Acad Sci U S A*. **106**, 3142–3147
38. Krois, A. S., Ferreon, J. C., Martinez-yamout, M. A., Dyson, H. J., and Wright, P. E. (2016) Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein. *Proc. Natl. Acad. Sci.* **113**, E1853–E1862
39. Krauskopf, K., Gebel, J., Kazemi, S., Tuppi, M., Löhr, F., Schäfer, B., Koch, J., Güntert, P.,

Dötsch, V., and Kehrloesser, S. (2018) Regulation of the Activity in the p53 Family Depends on the Organization of the Transactivation Domain. *Structure*. **26**, 1091-1100.e4

40. Dames, S. a, Martinez-Yamout, M., De Guzman, R. N., Dyson, H. J., and Wright, P. E. (2002) Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5271–5276
41. Feng, H., Jenkins, L. M. M., Durell, S. R., Hayashi, R., Mazur, S. J., Cherry, S., Tropea, J. E., Miller, M., Wlodawer, A., Appella, E., and Bai, Y. (2009) Structural basis for p300 Taz2-p53 TAD1 binding and modulation by phosphorylation. *Structure*. **17**, 202–210
42. Ferreon, J. C., Martinez-Yamout, M. a, Dyson, H. J., and Wright, P. E. (2009) Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13260–13265
43. Jenkins, L. M. M., Yamaguchi, H., Hayashi, R., Cherry, S., Tropea, J. E., Miller, M., Wlodawer, A., Appella, E., and Mazur, S. J. (2009) Two Distinct Motifs within the p53 Transactivation Domain Bind to the Taz2 Domain of p300 and Are Differentially Affected by Phosphorylation †. *Biochemistry*. **48**, 1244–1255
44. Miller Jenkins, L. M., Feng, H., Durell, S. R., Tagad, H. D., Mazur, S. J., Tropea, J. E., Bai, Y., and Appella, E. (2015) Characterization of the p300 Taz2-p53 TAD2 Complex and Comparison with the p300 Taz2-p53 TAD1 Complex. *Biochemistry*. **54**, 2001–2010
45. Oka, O., Waters, L. C., Strong, S. L., Dosanjh, N. S., Veverka, V., Muskett, F. W., Renshaw, P. S., Klempnauer, K.-H. H., and Carr, M. D. (2012) Interaction of the transactivation domain of B-Myb with the TAZ2 domain of the coactivator p300: molecular features and properties of the complex. *PLoS One*. **7**, e52906
46. Wang, F., Marshall, C. B., Yamamoto, K., Li, G.-Y. Y., Gasmi-Seabrook, G. M. C., Okada, H., Mak, T. W., and Ikura, M. (2012) Structures of KIX domain of CBP in complex with two FOXO3a transactivation domains reveal promiscuity and plasticity in coactivator recruitment. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6078–6083
47. Wojciak, J. M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009) Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *EMBO J.* **28**, 948–958
48. Ferreon, J. C., Lee, C. W., Arai, M., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2009) Cooperative regulation of p53 by modulation of ternary complex formation with CBP/p300 and HDM2. *Proc. Natl. Acad. Sci.* **106**, 6591–6596
49. Teufel, D. P., Freund, S. M., Bycroft, M., and Fersht, A. R. (2007) Four domains of p300 each bind tightly to a sequence spanning both transactivation subdomains of p53. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 7009–7014
50. Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795–2804
51. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
52. Guntert, P., and Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378
53. Guntert, P., and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR*. **62**, 453–457
54. De Guzman, R. N., Liu, H. Y., Martinez-Yamout, M., Dyson, H. J., and Wright, P. E. (2000) Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP. *J. Mol. Biol.* **303**, 243–253
55. Yi, P., Wang, Z., Feng, Q., Pintilie, G. D., Foulds, C. E., Lanz, R. B., Ludtke, S. J., Schmid, M. F.,

Chiu, W., and O'Malley, B. W. (2015) Structure of a Biologically Active Estrogen Receptor-Coactivator Complex on DNA. *Mol. Cell*. **57**, 1047–1058
56. Aspland, S. E., Bendall, H. H., and Murre, C. (2001) The role of E2A-PBX1 in leukemogenesis. *Oncogene*. **20**, 5708–5717
57. Sepp, M., Kannike, K., Eesmaa, A., Urb, M., and Timmusk, T. (2011) Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PLoS One*. **6**, e22138
58. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197–208
59. Park, S., Chen, W., Cierpicki, T., Tonelli, M., Cai, X., Speck, N. A., and Bushweller, J. H. (2009) Structure of the AML1-ETO eTAFH domain-HEB peptide complex and its contribution to AML1-ETO activity. *Blood*. **113**, 3558–3567
60. Zhao, F., McCarrick-Walmsley, R., Akerblad, P., Sigvardsson, M., and Kadesch, T. (2003) Inhibition of p300/CBP by early B-cell factor. *Mol. Cell. Biol.* **23**, 3837–3846
61. Lee, C. W., Ferreon, J. C., Ferreon, A. C. M., Arai, M., and Wright, P. E. (2010) Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 19290–19295
62. Kramer, A., and Schneider-Mergener, J. (1998) Synthesis and Screening of Peptide Libraries on Continuous Cellulose Membrane Supports. in *Combinatorial Peptide Library Protocols*, pp. 25–40, Humana Press, New Jersey, **87**, 25–40
63. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*. **6**, 277–293
64. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. **59**, 687–696
65. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*. **44**, 213–223
66. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol*. **231**, 1049–1067
67. Huang, Y. J., Powers, R., and Montelione, G. T. (2005) Protein NMR Recall, Precision, and F-measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics. *J. Am. Chem. Soc.* **127**, 1665–1674
68. Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating Protein Structures Determined by Structural Genomics Consortia Tools for Structure Quality Evaluation. *Proteins*. **795**, 778–795

FOOTNOTES

This work was supported by a Leukemia and Lymphoma Society operating grant to D.P.L. and Canada Canadian Institutes of Health Research operating grants to D.P.L (MOP 142694) and S.P.S. (MOP 106471) D.N.L. was supported by a Canadian Institutes of Health Research Postdoctoral Fellowship. M.R.L. and A.C.K. were recipients of Natural Science and Engineering Research Council of Canada CGS-M awards. The abbreviations used are: activation domain (AD), basic helix-loop-helix (bHLH), dissociation constant (K_d), B1 domain of protein G (GB1); isopropyl β-D-1-thiogalactopyranoside (IPTG), nuclear magnetic resonance (NMR), p300/CBP and ETO target in E proteins (PCET)

Table 1. Structural statistics of the E2A-AD1:TAZ2 complex.

Distance restraints used for structure calculations	**AD1**	**TAZ2**
Intra Residue	172	621
Sequential (l i-j l = 1)	290	722
Medium Range (1< l i-j l <5)	222	774
Long Range (l i-j l≥5)	1	409
Intermolecular	197	
Total	3408	
Hydrogen bond restraints	10	44
Dihedral Restraints		
φ	14	67
ψ	14	68
Energies (kcal mol^{-1})		
Total	-3328 ± 93	
Van der Waals	-758 ± 12	
RMSD from experimental restraints		
Distance constraints (Å)	0.021 ± 0.002	
Dihedral angles (°)	0.418 ± 0.088	
RMSD from idealized geometry		
Bond lengths (Å)	0.015 ± 0.001	
Bond angles (°)	0.973 ± 0.013	
Global quality scores (Raw/Z-score)		
Verify3D	0.33/-2.09	
Procheck (phi-psi)	-0.17/-0.35	
Procheck (all)	-0.33/-1.95	
MolProbity clash score	16.88/-1.37	
RPF Validation		
Recall	0.906	
Precision	0.872	
F-measure	0.889	
Discriminating power	0.807	
Ramachandran statistics (%)[1]		
Residues in most favoured regions	94.4	
Residues in additionally allowed regions	4.6	
Residues in generously allowed regions	0.2	
Residues in disallowed regions	0.9	
RMSD to mean structure (Å)[1]		
Backbone atoms	0.51 ± 0.06	
Heavy atoms	0.83 ± 0.06	

[1] Using converged regions, E2A-AD1 residues Asp13-Pro29 and TAZ2 residues Ser1726-Lys1812.

FIGURES

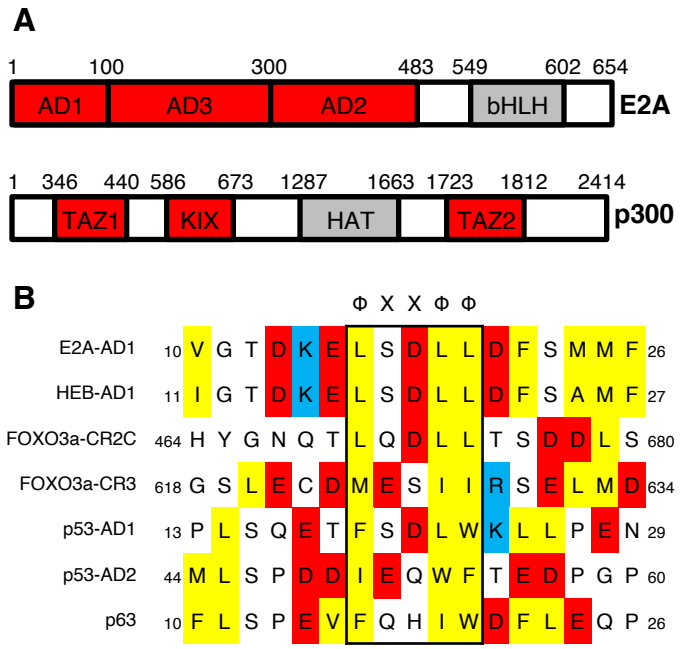


Figure 1. Domain architecture of E2A and CBP/p300**.** (A) Schematic representation of E2A and the human histone acetyltransferase (HAT) p300, with the relevant domains (E2A: activation domains AD1, AD2, and AD3 and the C-terminal DNA-binding basic helix-loop-helix (bHLH) domain; p300: The TAZ1, KIX, and TAZ2 protein-protein interaction domains and the HAT catalytic domain) indicated. (B) Sequence alignment of the ϕ-x-x-ϕ-ϕ motifs of the activation domains (AD) of E2A, HEB, FOXO3a, p53, and p63, where ϕ represents a hydrophobic amino acid residue and x any amino acid residue. Hydrophobic, acidic, and basic amino acid residues are coloured yellow, red, and blue, respectively. Numbering is in accordance with the native protein sequence.

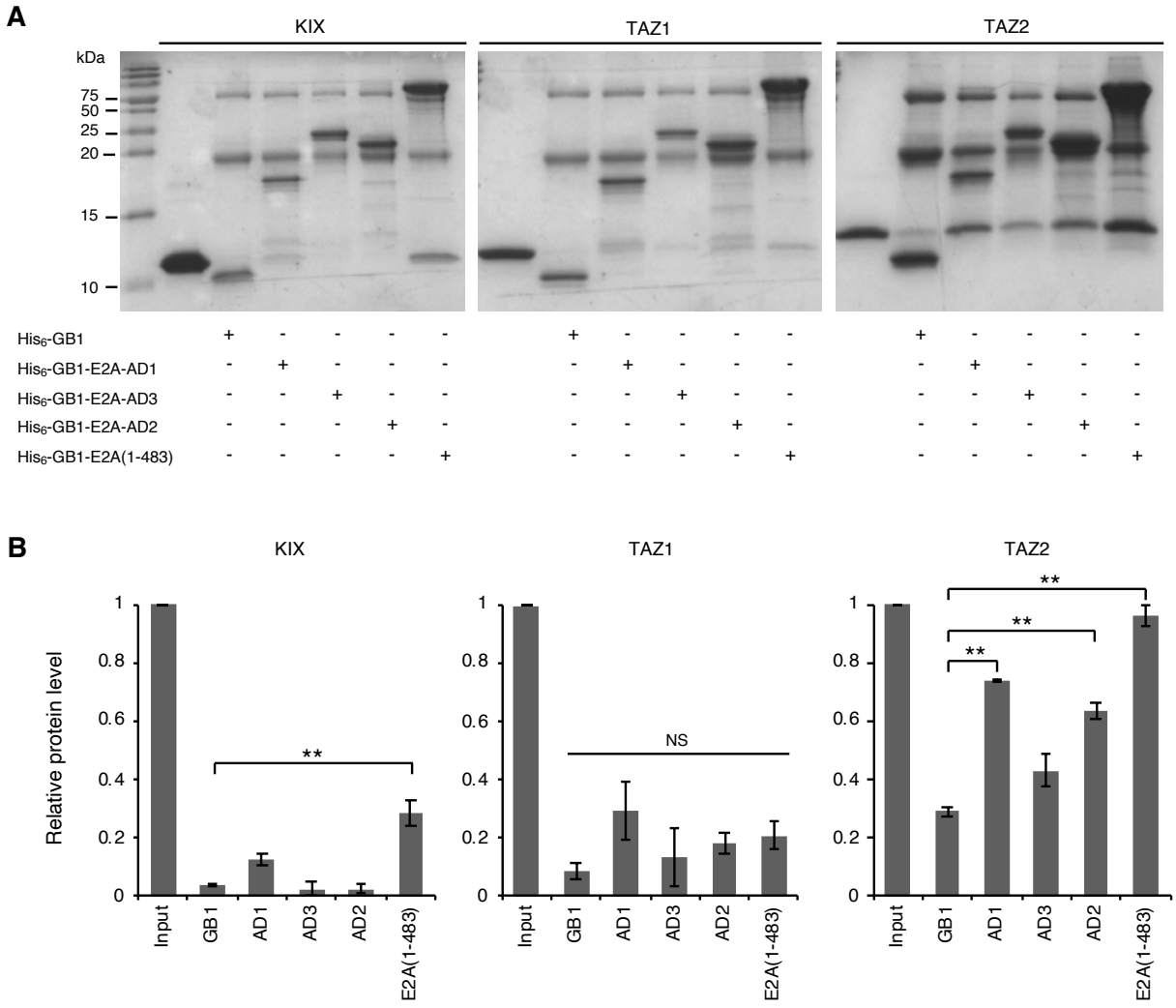


Figure 2. E2A-AD1 binds the TAZ2 domain of CBP/p300. (A) 15% SDS-PAGE analysis of the ability of various purified His_6-GB1-E2A fragments immobilized on IgG agarose to pulldown purified recombinant KIX, TAZ1, and TAZ2 domains of p300. The left lanes of each panel represent controls illustrating the migration of the recombinant isolated p300 domains. (B) Quantitative analysis of the His_6-GB1-E2A pulldowns presented in (A). Significance between pulldown protein levels of TAZ1, TAZ2 and KIX by GB1 (negative control) and GB1-E2A constructs is indicated by brackets. Error bars represent the standard deviation (** $p < 0.01$).

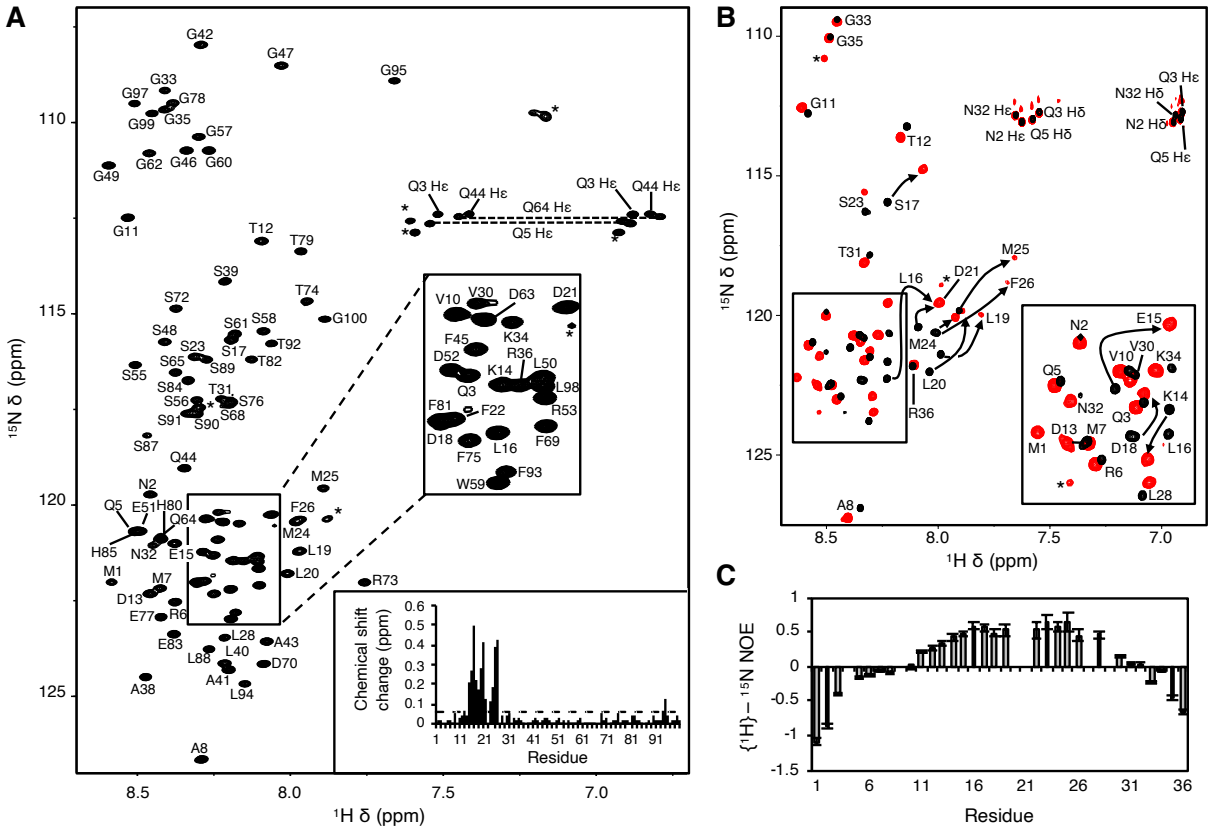


Figure 3. The ϕ-x-x-ϕ-ϕ containing region of E2A-AD1(1-37) adopts an ordered structure upon TAZ2 binding. (A) The ^{1}H-^{15}N HSQC of uniformly ^{15}N-labeled E2A-AD1. Resonance assignments are indicated and * denotes resonances that are due to impurities or could not be assigned due to spectral overlap. The inset indicates the weighted chemical shift change ($\Delta\delta = [(0.17\Delta\delta_N)^2 + (\Delta\delta_{HN})^2]^{1/2}$ that each residue of E2A-AD1 experiences upon addition of saturating amounts of TAZ2. (B) Overlay of ^{1}H-^{15}N HSQC spectra of uniformly ^{15}N-labeled E2A-AD1(1-37) in the absence (black) and presence (red) of saturating amounts of TAZ2. Resonance assignments are indicated, and arrowed lines indicate the directional movement of resonances upon addition of TAZ2. Resonances attributed to impurities are indicated with *. (C) A plot of steady state {^{1}H}-^{15}N NOE values recorded at 14.1 T as a function of E2A-AD1(1-37) residue number when bound to TAZ2.

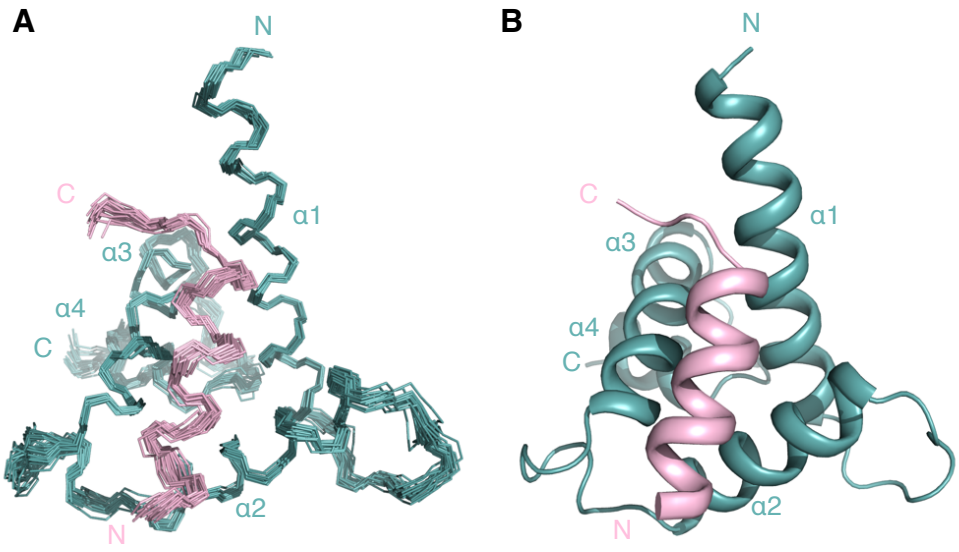


Figure 4. Structure of the E2A-AD1(1-37):TAZ2 complex. (A) Backbone superposition of the 20 lowest-energy structures of the E2A-AD1(1-37):TAZ2 complex, with the backbone atoms (N, C^{α} and C') of Asp13-Pro29 of E2A-AD1(1-37) (pink) and Ser1726-Lys1812 of the TAZ2 domain (teal) displayed. (B) Backbone ribbon illustration of the lowest energy structure of the E2A-AD1(1-37):TAZ2 complex. The helices of the TAZ2 domain are labeled α1, α2, α3, and α4, and the N- and C-termini of E2A-AD1(1-37) and TAZ2 are indicated.

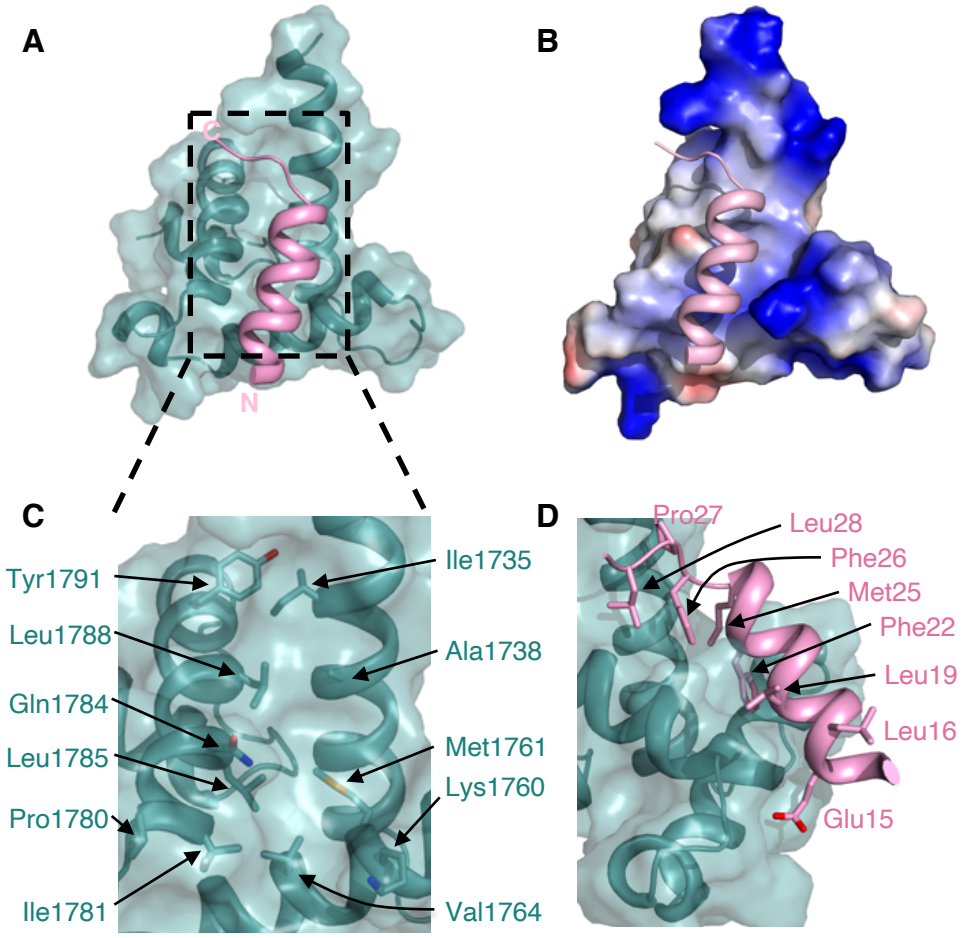


Figure 5. The E2A-AD1(1-37):TAZ2 interface. (A) Backbone ribbon representation of the lowest energy structure of the E2A-AD1(1-37) (pink):TAZ2 (teal) complex onto which a transparent surface representation of the TAZ2 domain also displayed. (B) Electrostatic surface representation of the E2A-AD1 binding face of TAZ2 domain, where blue and red shadings depict positively and negatively charged regions, respectively. (C) Expanded view of transparent surface of the TAZ2 domain displaying the backbone ribbon and residues involved in intermolecular van der Waals contacts with E2A-AD1(1-37). (D) Backbone ribbon representation of E2A-AD1(1-37) (pink) showing those residues making intermolecular van der Waals contacts on the surface of the TAZ2 domain.

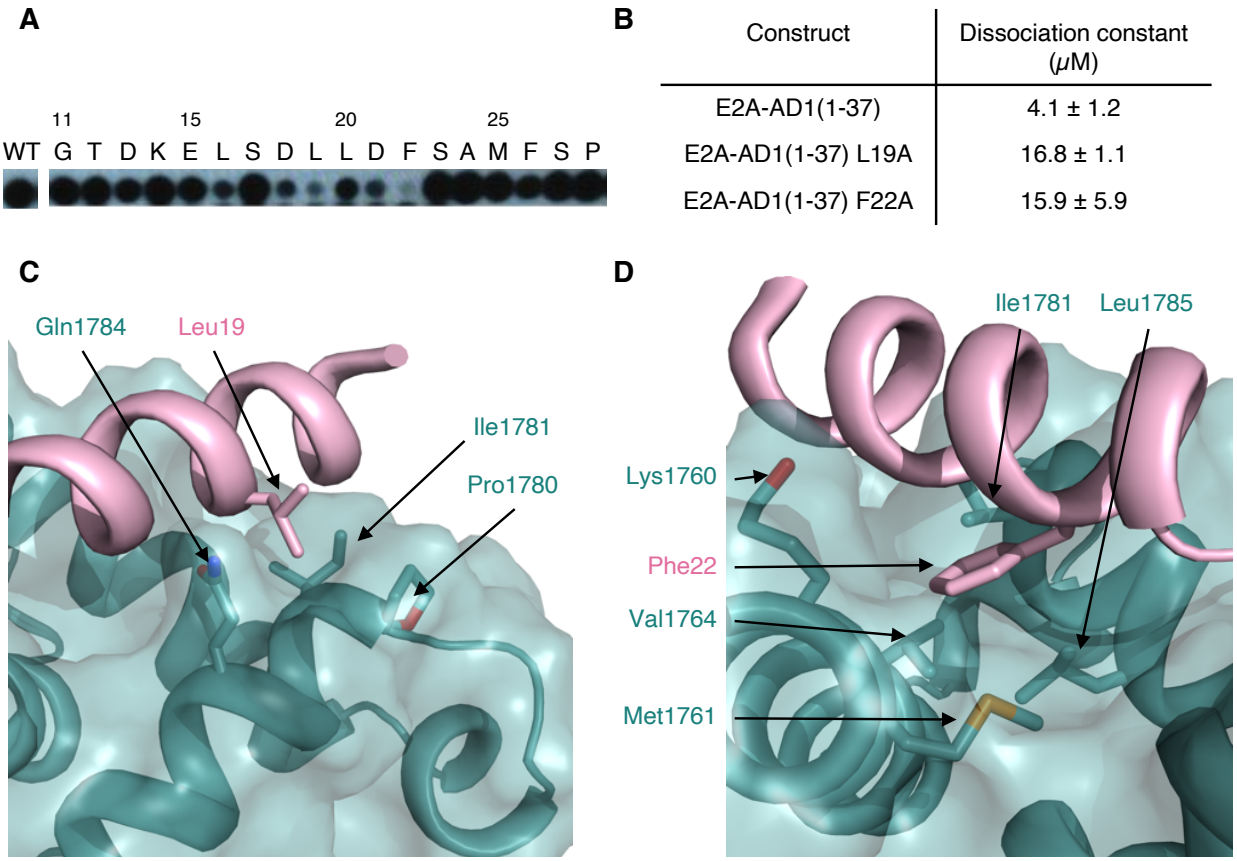


Figure 6. E2A-AD1 mutants impair TAZ2 binding. (A) Membranes spotted with HEB-AD1(11-28) peptides in which each residue had been individually substituted with alanine were probed with His_6-GB1-TAZ2. The sequence of HEB is indicated using E2A numbering and each position of HEB-AD1(11-28) was mutated to alanine and spotted onto an array. The signal present from TAZ2 binding to each alanine mutant of HEB-AD1(11-28) is indicated as a spot below the sequence. (B) The isothermal titration calorimetric dissociation constants determined for the interaction between wild-type, Leu19Ala and Phe22Ala E2A-AD1(1-37) and the isolated TAZ2 domain. (C,D) Ribbon representation of the E2A-AD1(1-37):TAZ2 complex with a transparent TAZ2 surface showing the position of Leu19 (C) and Phe22 (D) of E2A-AD1 (pink) on the TAZ2 (teal) surface.

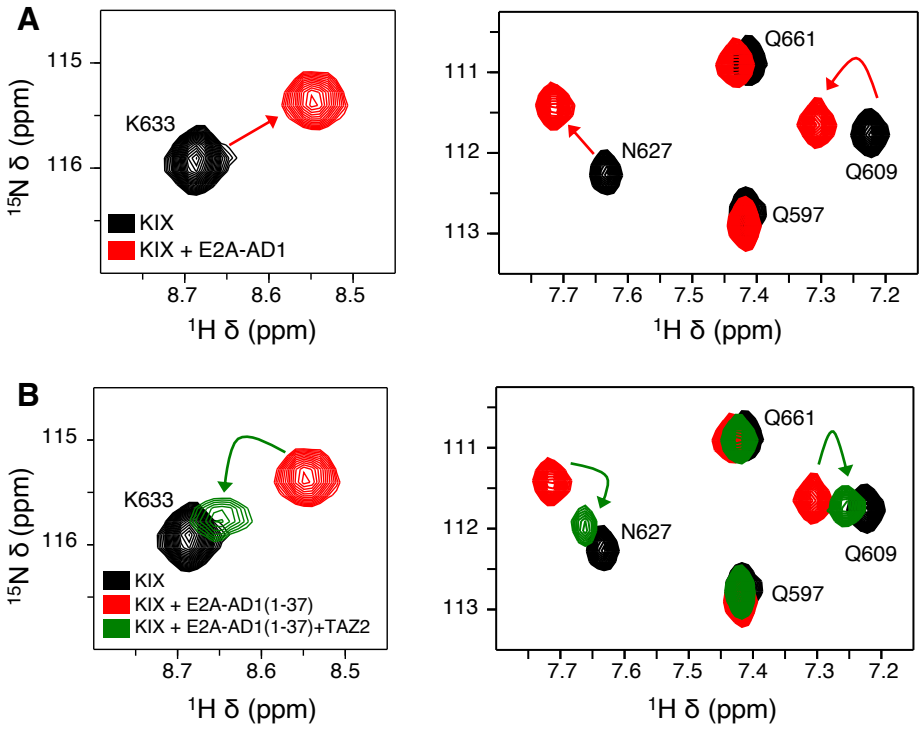


Figure 7. TAZ2 and KIX domains compete for E2A-AD1 *in vitro*. (A) Overlay of two selected regions from 2D ^{1}H-^{15}N HSQC spectra of 200 μM ^{15}N-KIX in the absence (black) and presence (red) of 300 μM unlabeled E2A-AD1(1-37). (B) Overlay of the same two regions depicted in panel A after the addition of 500 μM unlabeled TAZ2. The direction of backbone amide resonance change is depicted by red and green arrows in panels A and B, respectively.

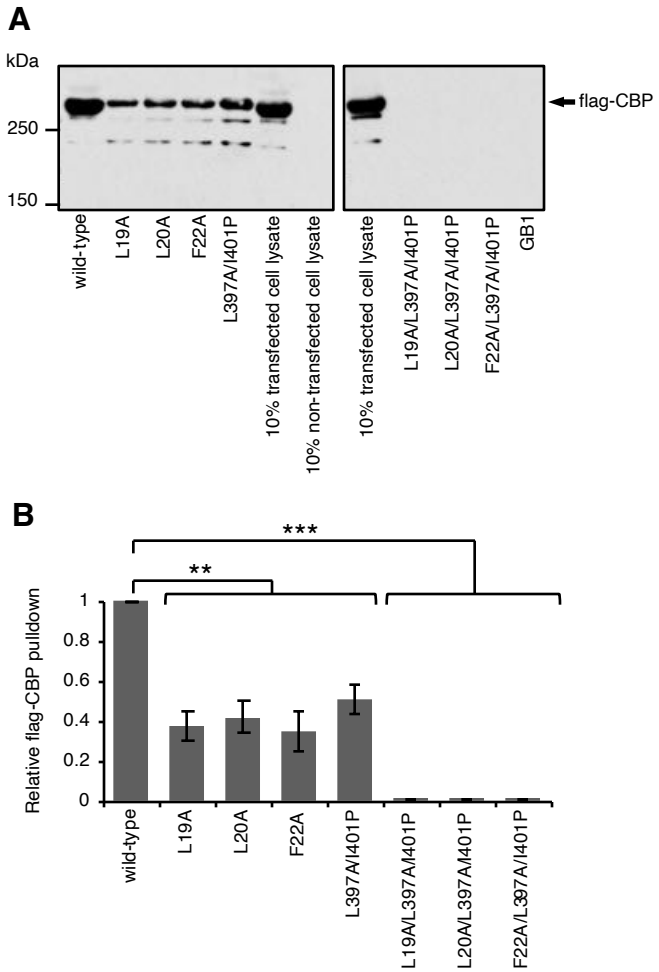


Figure 8. Mutation of AD1 or AD2 of E2A(1-483) reduces its ability to pulldown CBP. (A) Immunoblotting of comparative pull-downs of flag-CBP by purified His6-GB1-E2A(1-483) constructs immobilized on IgG agarose. (B) Relative quantitative analysis of the western blots performed in triplicate. ImageJ was used to measure the pixel density for each lane. The y-axis values represent the ratio of flag-CBP pull-down by His6-GB1-E2A(1-483) mutants to flag-CBP pull-down by His6-GB1-E2A wild-type. Error bars represent standard deviation (n=3, ** $p < 0.0005$, *** $p < 0.0001$).

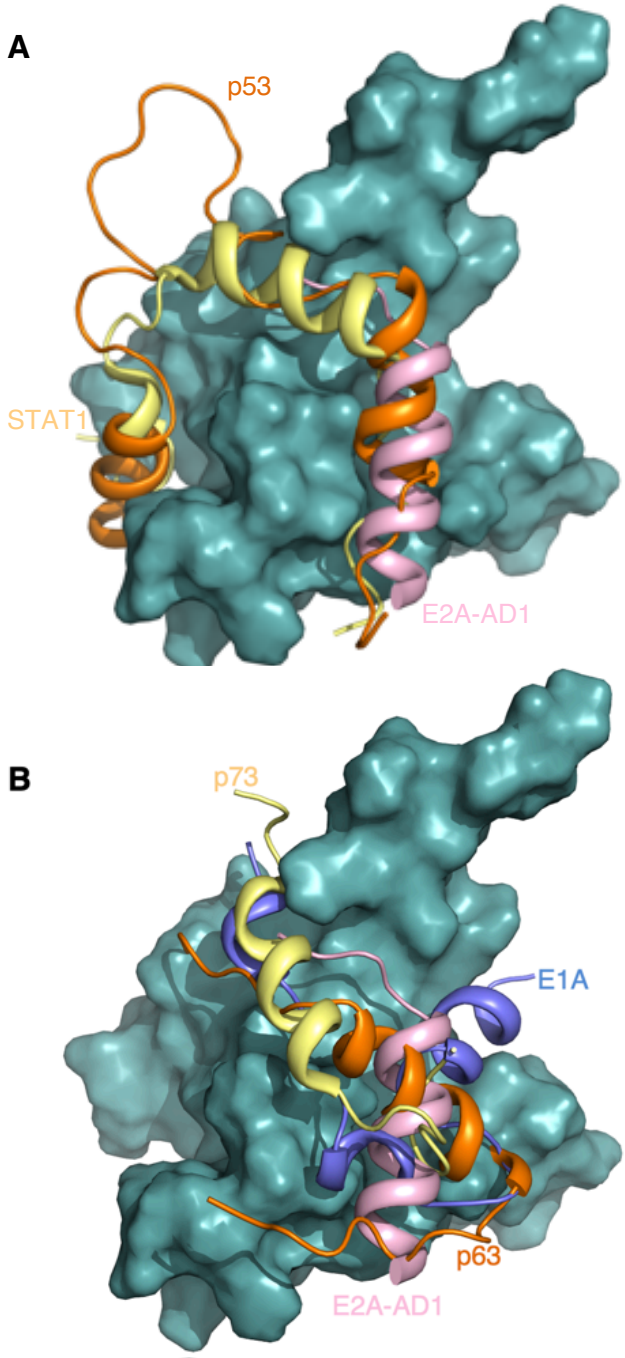
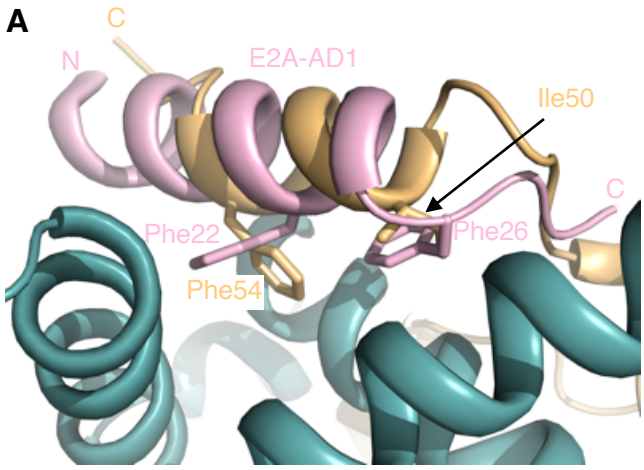


Figure 9. Transcription factors adopt diverse structures and orientations when interacting with Taz2. (A) Backbone ribbon representation of p53 (5HPD; orange (38)), E2A-AD1(1-37) (pink), and STAT1 (2KA6; yellow (47)) in complex with TAZ2, which is shown as a teal surface. In (B) the binding orientations of E2A-AD1(1-37) (pink), E1A (2KJE; blue (42)), p63 (6FGN; orange), and p73 (6FGS; yellow (39)) are shown.



B

E2A	18 -D	L	L	D	F	S	M	M	F	- 26
p53	58 -P	D	E	T	F	W	Q	E	I	- 50

Reversed

Figure 10. E2A-AD1 and p53 bind TAZ2 with opposite orientations. (A) Ribbon representation of a superposition of the E2A-AD1(1-37):TAZ2 and p53:TAZ2 (38) complexes, with only TAZ2 (teal), E2A-AD1(1-37) (pink), and p53 (yellow) shown. The side chains of Phe22 and Phe26 of E2A-AD1 as well as Ile50 and the Phe54 of p53 are displayed as sticks. (B) A sequence alignment of E2A-AD1 and p53. The sequence of p53 is reversed with respect to E2A-AD1 and the residues highlighted in (a) are coloured red.