Published in final edited form as:

Biochemistry. 2021 December 28; 60(51): 3887–3898. doi:10.1021/acs.biochem.1c00669.

Characterization of the high-affinity fuzzy complex between the disordered domain of the E7 oncoprotein from high-risk HPV and the TAZ2 domain of CBP

Michael W. Risør[1,3], **Ariane L. Jansma**[2,3], **Natasha Medici**[2], **Brittany Thomas**[2], **H. Jane Dyson**[1,*], **Peter E. Wright**[1,*]

[1]Department of Integrative Structural and Computational Biology and Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California, 92037, U.S.A.

[2]Department of Chemistry, Point Loma Nazarene University, San Diego, California, 92106, U.S.A.

[3]Joint first author

Abstract

The intrinsically disordered N-terminal region of the E7 protein from high-risk human papillomavirus (HPV) strains is responsible for oncogenic transformation of host cells through its interaction with a number of cellular factors, including the TAZ2 domain of the transcriptional coactivator CREB-binding protein (CBP). Using a variety of spectroscopic and biochemical tools we find that, despite its nanomolar affinity, the HPV16 E7 complex with TAZ2 is disordered and highly dynamic. The HPV16 E7 protein does not adopt a single conformation on the surface of TAZ2 but engages promiscuously with its target through multiple interactions involving two conserved motifs, termed CR1 and CR2, that occupy an extensive binding surface on TAZ2. The fuzzy nature of the complex is a reflection of the promiscuous binding repertoire of viral proteins, which must efficiently dysregulate host cell processes by binding to a variety of host factors in the cellular environment.

Graphical Abstract

[*]Author for correspondence: H. Jane Dyson, Phone: 1-858-784-2223, dyson@scripps.edu, Peter E. Wright, Phone: 1-858-784-9721, wright@scripps.edu.

AUTHOR CONTRIBUTIONS

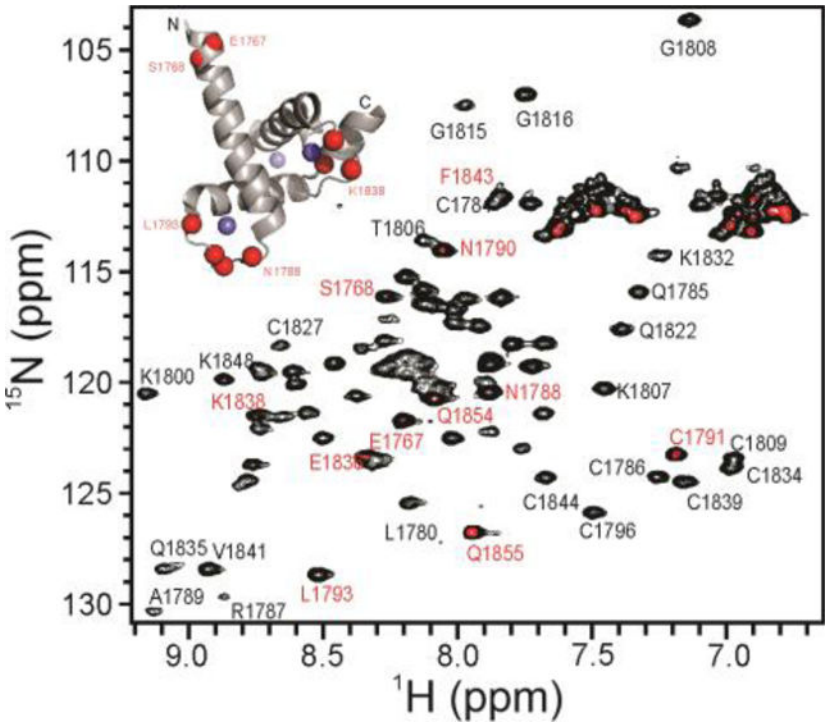
PEW, HJD, AJ and MWR planned experiments; AJ, MWR, NM and BT performed experiments; PEW, HJD, AJ and MWR analyzed data; AJ, MWR, HJD and PEW wrote the paper.

ACCESSION NUMBERS

E7 protein from human papillomavirus (strain 16) UniProt P03129.

SUPPORTING INFORMATION

Full NMR spectra, NOESY spectra and table showing intermolecular NOEs, and SAXS data.



Keywords

human papillomavirus; intrinsically disordered protein; transcription factor; viral oncoprotein

INTRODUCTION

Intrinsic disorder in proteins is critical for a wide variety of cellular processes.[1–5] Intrinsically disordered proteins (IDPs) consist mainly, if not entirely, of polypeptide chains that do not have stable 3D structures, while intrinsically disordered regions (IDRs) constitute the disordered segments of proteins that might otherwise contain structured regions. The dynamic nature of disordered proteins allows for a range of options for complex formation: they may fold into an ordered structure upon binding, or a complex may be formed that retains various levels of disorder.[6–9] This versatility is particularly useful to viruses, which generally utilize IDPs and IDRs to a greater extent than their host systems, owing to their need to maximize functional diversity with minimal genetic information.[10–13]

Human papillomavirus (HPV) is a family of DNA viruses with more than 200 subtypes with varying propensities for causing cancer.[14] Infection with most subtypes, classified as “low-risk”, results in the formation of benign skin lesions, but infection with “high-risk” HPV leads to oncogenic transformation of host cells.[15–17] High-risk HPV serotypes are responsible for virtually all instances of cervical, anal, rectal, and penile cancers as well as an increasingly high amount of oropharyngeal cancers.[16] Of the high-risk serotypes, HPV16 is particularly oncogenic and responsible for more than 50% of all invasive cervical cancers.[18] Tumor growth is induced through the activity of the two major oncoproteins E6 and E7, which act synergistically to transform and immortalize epithelial cells.[19,20] E7

dysregulates the cell cycle to advance cellular and viral replication while E6 simultaneously inhibits host cell apoptosis.[21–23] Both E6 and E7 employ IDRs to achieve their various functions and for E7 this feature is essential for its transforming activity. The E7 protein is a relatively small (98 amino acids for HPV16), highly acidic protein consisting of three conserved regions denoted CR1, CR2 and CR3 (Figure 1A). CR1 and CR2 are intrinsically disordered and represent about half of the protein;[10,24–26] CR3 is a structured zinc-binding domain responsible for E7 homodimer formation.[24,27–29] Additionally, this protein shows almost complete sequence conservation in highly invasive cancer isolates across the globe.[30] Despite its small size and sequence conservation, E7 has a highly promiscuous target binding activity[12,31] and interacts with a myriad of host cell factors to target cellular processes such as transcriptional and immune regulation, cell signaling, and protein homeostasis.[31–35]

The best-characterized target of E7 is the retinoblastoma tumor suppressor protein (pRb),[20] which regulates cell-cycle progression. The acetylation and degradation of pRb is mediated by HPV16 E7 through the recruitment of the cyclic-AMP response element binding protein (CBP) into a ternary complex.[37] CBP and its paralog p300 (Figure 1B) are multi-domain transcriptional coactivators responsible for the activation of numerous pathways and are key regulators of cell growth and differentiation.[38,39] Because of their central role in cellular regulation, these proteins are frequently targeted by viral oncoproteins.[40] The HPV E7 protein interacts with the TAZ2 domain of CBP through its CR1 and CR2 motifs (Figure 1C).[37] The TAZ2 domain also acts as the binding site for the adenoviral E1A protein[41] as well as numerous cellular IDRs.[38]

In the present work, we undertook detailed characterization of the complex between a 51-residue peptide encompassing the intrinsically disordered N-terminal region of E7 (Figure 1A) and the TAZ2 domain of CBP/p300. The intent was to solve the high-resolution structure for this complex. Crystals were never obtained for the complex, so structure determination was pursued by solution NMR spectroscopy. Intriguingly, the interaction between E7 and TAZ2 is highly dynamic despite a dissociation constant in the nanomolar range, and the NMR data indicate the presence of multiple co-existing conformations. Truncated and mutated versions of E7 were employed to decipher the contributions from the CR1 and CR2 domains individually. It appears that while CR2 has a slightly higher affinity for TAZ2, the fuzzy complex formed by the disordered N-terminal region of E7 samples conformations where CR1 and CR2 both bind to the same general surface of TAZ2, which is large enough to accommodate both motifs simultaneously. This detailed insight allows us to present a thorough characterization of the structural properties of this highly complicated system.

MATERIALS AND METHODS

Protein expression and synthetic peptides

TAZ2 (residues 1764–1855) domains of wild-type mouse CBP were expressed as previously described.[36,42] The E7 sequences correspond to the human papilloma virus strain HPV16. All of the TAZ2 interacting partners were expressed as His_6-GB1 fusions, including E7(1–51), E7(17–51), the N-terminal fusions E7(1–51)-TAZ2, E7(1–51)-GS_3-TAZ2, and

E7(1–40)-TAZ2, the C-terminal fusions TAZ2-E7(1–51) and TAZ2-GS_2-E7(1–51), as well as E7(1–51) C24A/A50C for labeling with Alexa 594 for fluorescence anisotropy measurements, and E7(1–51)H9C/C24A for incorporation of a nitroxide spin label. All constructs were expressed for 16–20 hours at 16°C in BL21DE3 (+DNAY) cells. After growth at 37°C to OD ~ 0.8 cells were induced by addition of 0.5 mM IPTG. Non-labeled samples were prepared by growth in LB or M9 medium, and uniformly isotope labeled samples were prepared by growth in M9 medium supplemented with 1 g/L $^{15}NH_4Cl$ and 2 g/L ^{13}C-D-glucose. The growth medium was supplemented with 150 μM $ZnSO_4$ for all constructs containing TAZ2. E7 CR1 (residues 1–18; NH_2-MHGDTPTLHEYMLDLQPE-NH_2) and CR2 (residues 18–37; NH_2-ETTDLYCYEQLNDSSEEEDE) peptides were ordered as TFA salts from Biomatik with a purity above 95%. The peptides were resuspended in 50 mM Tris buffer, pH 8.0 and buffer-exchanged into assay buffer or NMR buffer.

Protein Purification

TAZ2 and E7(1–51) constructs were purified as previously described.[36,37,42] Final purity and label incorporation were confirmed by mass spectrometry. All other bacterial pellets from E7 peptides and fusions were lysed in pH 7.2 buffer containing 20 mM Tris, 200 mM NaCl, and 1 mM DTT, and purified by application of the lysate to a Complete Ni-resin. After washing and elution by 250 mM imidazole, the samples were dialyzed into low-salt buffer with overnight TEV cleavage at a 1:400 molar ratio; cleavage was confirmed by analytical HPLC. For E7(1–51) and E7(17–51), the low salt buffer was 20 mM Tris, 50 mM NaCl, pH 7.2, and further purification was carried out on a Hitrap Q column with NaCl gradient elution. Purified peptides were dialyzed into either NMR buffer (20 mM Tris, 50 mM NaCl, 1 mM DTT, pH 6.8) or assay buffer (20 mM Tris, 50 mM NaCl, 1 mM DTT, pH 7.0). For E7-TAZ fusions, the low salt buffer was 20 mM MES, 50 mM NaCl, pH 6.3, and further purification was carried out on a Hitrap SP column with NaCl gradient elution. Purified fusions were dialyzed into NMR buffer and concentrated.

Phosphorylation and D_2O samples

E7(1–51) was phosphorylated as previously described[37] and confirmed by mass spectrometry. E7(17–51) was phosphorylated at ~ 400 μM with home-made CKII for three hours at 30°C at a protein to kinase molar ratio of 400:1 in a pH 7.2 buffer containing 20 mM Tris, 100 mM NaCl, 10 mM $MgCl_2$, and 2.5 mM ATP. The reactions were followed by analytical HPLC to ensure completion of the phosphorylation reaction. The phosphorylated peptides were loaded onto a Hitrap Q column and eluted with a NaCl gradient. The peptides were dialyzed into NMR buffer and concentrated. D_2O NMR samples for ppE7(1–51) with TAZ2 were obtained by dialyzing the mixed complex into D_2O buffer with 20 mM d-Tris, 50 mM NaCl, 2 mM d-DTT at pD 6.4 (pH 6.8). D_2O NMR samples for ppE7(17–51) and TAZ2 were obtained by exchanging individual components into D_2O NMR buffer by a NAP-5 column before mixing. Formation of the TAZ2:ppE7(17–51) complex was subject to extensive aggregation, and the TAZ:E7 ratio in the final soluble sample was estimated from analytical HPLC peak integrals.

Fluorescence anisotropy

E7(1–51) C24A/A50C was labeled with Alexa fluor 594 dye using a 2.5 molar excess of the reactive dye in a pH 7.2 buffer for two hours at room temperature. The labeled protein was separated from the dye and exchanged into assay buffer by a NAP-5 desalting column. Ensemble fluorescence anisotropy measurements were carried out on a Fluorolog®−3 instrument using extinction and emission wavelengths of 590 and 617 nm, respectively, with a concentration of 20 nM of A594-E7(1–51). Direct TAZ2 titrations were carried to complete saturation of binding (1.4 μM). For competition experiments with E7 peptides, a sample of A594-E7(1–51) with 0.8 μM TAZ2 was prepared (> 90% saturation), and the decreasing fluorescence anisotropy was measured as a function of the competing peptide concentration. The direct titration series was fitted using a standard 1-site binding equilibrium[43]:

$$A_T = A_{free} + \left(A_{bound} - A_{free}\right) \cdot \frac{(L1_T + K_d + R_T) - \sqrt{(L1_T - K_d - R_T)^2 - 4L1_T R_T}}{2}$$

Where A_T is the measured fluorescence anisotropy, A_{free} is the anisotropy without TAZ2, and A_{bound} is the saturated signal upon full TAZ2 binding, $L1_T$ is the cuvette concentration of A594-E7(1–51), and R_T is the TAZ2 concentration. The competition experiments were fitted as previously described[44]:

$$A_T = \left(\gamma A_{bound} + A_{free}\right)/\left(1 + \gamma\right)$$

where γ is the fraction of L1 peptide bound, $\gamma = [RL1]/([L1_T] - [RL1])$. Because γ decreases as a function of the titrant concentration of the competing peptide we can derive the apparent K_d for the competing peptide. For the full expression, see previous work.[44]

Spin label incorporation

4-maleimido-TEMPO (Sigma-Aldrich) was added in 3 molar excess to ppE7(1–51)H9C,C24A in 25 mM Tris buffer at pH 7 and allowed to incubate for four hours. The protein was purified by HPLC.

NMR spectroscopy

NMR experiments were conducted at 30°C on various Bruker spectrometers (Avance600, Avance700 and Avance800 equipped with cryoprobes, and Avance900). Backbone and side chain assignments for ^{15}N,^{13}C-labeled E7 constructs alone and in complex with TAZ2 were obtained using ^{1}H-^{15}N-HSQC, ^{1}H-^{13}C-HSQC, 3D-HNCA, 3D-HNCACB, 3D-HNCO, HCCH-TOCSY, and HCCH-COSY spectra. TAZ2 methyl assignments could be made through the recorded 3D experiments, but several CH_2 groups and Hα protons could not be unambiguously assigned. For titrations of E7 peptides with ^{15}N-TAZ2 or titrations of TAZ2 with ^{15}N-E7 peptides, we carried out standard ^{1}H-^{15}N HSQC experiments on the unbound forms and on the samples after addition of either TAZ2 or E7 peptides to the specified concentration ratios. Experiments were conducted at 30°C and 2 mM fresh DTT was added to each sample before transfer to the NMR tube. The concentration of the labeled

components was between 50 and 100 μM. The (^{15}N,^{13}C)-TAZ2:E7(1–51), (^{15}N,^{13}C)-TAZ2:E7(17–51), (^{15}N,^{13}C)-E7(1–51):TAZ2, and (^{15}N,^{13}C)-E7(17–51):TAZ2 complexes in D_2O were used to record edit-filter and filter-edit NOESY-HSQC experiments[45] to detect intermolecular NOE signals; mixing times were 200ms. Regular ^{13}C-NOESY-HSQC spectra were also recorded, with mixing times of 80 ms. A ^{15}N-NOESY-HSQC spectrum was also recorded for the (^{15}N,^{13}C)-E7(17–51):TAZ2 complex in H_2O NMR buffer.

Data processing

NMR spectra were processed and analyzed using NMRPipe, CCPN, and SPARKY, and referenced to DSS at 0 ppm. Chemical shift perturbations were calculated as $\Delta\delta_{ave} = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{1/2}$. Dihedral angles were calculated using the TALOS+N server,[46] and secondary structural propensities were calculated using the POTENCI prediction method.[47] Dissociation constants from ^{15}N-TAZ2 NMR titrations were derived by global fitting to a 1-site binding model using the in-house program nmrKd.[48]

Small Angle X-ray scattering of E7(1–51) and generation of peptide models

Samples for SAXS were prepared at 4.5 mg/mL, 2.25 mg/mL, and 1.13 mg/mL for E7(1–51) in 20 mM Tris-HCl, 100 mM NaCl, pH 7.0, and 2 mM fresh DTT. Data were collected on the Bio-SAXS beam line BL4–2 at Stanford Synchrotron Radiation Light-source (SSRL) using a Pilatus 300K detector and a beam energy of 11 keV. All data were collected up to a maximum q of 0.74 $Å^{-1}$ (1.1 m sample-to-detector distance). Scattering images were recorded with 1 s exposures using the data acquisition program Blu-ICE.[49] The data processing program SasTool was used for scaling and azimuthal integration. A buffer scattering profile with 30 frames was collected before each sample and averaged and subtracted from subsequent images to produce the scattering curve for each sample frame. Further data analysis was performed using tools in the ATSAS package[50] or online programs mentioned below. Sample data frames unaffected by radiation damage were averaged in PRIMUS to produce the final sample scattering curves. Data were processed and scaled for all concentrations of E7(1–51). The experimental data were analyzed by the online EOM service[51] to compute the R_g distribution of the E7(1–51) conformer ensemble. Conformations of the E7(1–51) peptide in best agreement with the experimental SAXS data were computed with the Allosmod-FOXS online server.[52,53]

Structural models generated by Haddock

We used the online version of Haddock2.2[54] for docking of E7 peptides onto the surface of TAZ2. The TAZ2 starting structure was adopted from the solution structure of E1A in complex with TAZ2 (PDB ID 2KJE,[41] where helix 1 is slightly more open compared to the non-complexed TAZ2 structure (PDB ID 1F81,[36]). The E7(17–40) peptide starting structures were generated from the 20 best models of E7(1–51) based on the computed output from the Allosmod-FOXS server using the experimental SAXS data on E7(1–51) as input. The Haddock docking was run in Guru mode, and the resulting Haddock models are the result of three iterations: rigid body energy minimization (it0), semi-flexible refinement in torsion angle space (it1), and final refinement in explicit solvent (water). Full flexibility (torsion angles and side chain dynamics) was allowed for all E7 peptide variants through all iterations to account for their dynamic nature and the TAZ2 structure was allowed to

be semi-flexible (side chain reorientation). Several different simulations were initiated with different sets of constraints, mostly intermolecular.

RESULTS

NMR analysis of the HPV16 E7-TAZ2 complex

Previous NMR studies of the HPV E7-TAZ2 complex[37] have revealed behavior that differs from that of other complexes of structured CBP domains with intrinsically disordered partners.[40] The intrinsically disordered residues 1–51 of the E7 protein comprise the CR1 and CR2 domains (Figure 1A). CR2 includes the LXCXE motif, the primary binding site for pRb,[20] as well as the two serine residues, S31 and S32, that are phosphorylated in cells by casein kinase 2 (CK2). The primary binding domain for TAZ2 includes residues in both the CR1 and CR2 domains (Figure 1A),[37] and the binding affinity for the doubly-phosphorylated peptide HPV16 ppE7(1–51) for TAZ2 is in the low nanomolar range, as measured by both fluorescence anisotropy and NMR spectroscopy.[37]

^{1}H-^{15}N heteronuclear single-quantum coherence (HSQC) spectra of ^{15}N-labeled ppE7(1–51) titrated with unlabeled TAZ2 show stepwise chemical shift perturbations as a function of increasing molar amounts of TAZ2 (Figure 2A), an indication that the complex and its components are in fast exchange on the NMR chemical shift timescale. This observation is unexpected in view of the high affinity of the complex, for which slow exchange might be expected. However, other systems have been documented where high affinity interactions exhibit fast exchange on the NMR time scale.[55]

Although the ^{1}H-^{15}N HSQC spectrum of ppE7(1–51) bound to TAZ2 is reasonably well dispersed, the ^{1}H-^{13}C HSQC spectrum is highly overlapped (Figure 2B). Of the methyl-containing residues in E7(1–51), only the 5 leucines and 2 of the 4 threonines form part of the TAZ2 binding site.[37] The three alanines and the single isoleucine are outside the binding site and experience little change in amide chemical shift in the presence of TAZ2; the alanine methyls would therefore be expected to have similar, if not identical, chemical shifts in the complex, as they do (Figure 2B). However, if ppE7(1–51) bound tightly to TAZ2 in a single conformation, we would expect the methyl resonances of the leucines and two of the threonines to be well-dispersed, as they would presumably be present in different chemical environments. Figure 2B shows that the methyl cross peaks of Thr and Leu are tightly clustered in the ^{1}H-^{13}C spectrum of the complex.

In an attempt to resolve some of the ambiguities in the methyl resonance assignments of ppE7(1–51) in the complex with TAZ2, we prepared a truncated form of the peptide, residues 17–51 (Figure 1A), removing the CR1 region and 3 of the 5 leucines. NMR analysis of the truncated E7(17–51) shows reduced spectral overlap and provides further information on the role of individual E7 sequence motifs in binding TAZ2. A comparison of the ^{1}H-^{15}N HSQC spectra of free ppE7(1–51) and ppE7(17–51) (Figure S1A) shows close correspondence between the cross peaks of residues 18–51 in the two peptides, indicating that the truncation does not alter the conformational ensemble of the free peptide. Addition of TAZ2 to ^{15}N-labeled ppE7(17–51) (Figure S1B) results in chemical shift changes for residues 20–29 that are very similar to those of the corresponding residues in ppE7(1–51)

(Figure 2A). The methyl resonances of ppE7(17–51) remain overlapped even in the presence of TAZ2 (Figure 3A). A further attempt to resolve the leucine methyls in ppE7(1–51) was made by substituting L15 with valine; the leucine methyls remain overlapped in this mutant (Figure 3B). NOESY spectra of the L15V mutant-TAZ2 complex were too poor to provide information on the local environment of L15 by comparison with the unchanged ppE7(1–51) complex. For both of these mutant forms [ppE7(17–51) and ppE7(1–51)(L15V)], the addition of TAZ2 results in only small chemical shift perturbations (e.g., ~0.1ppm in the ^{1}H dimension) and no increase in dispersion of the methyl cross peaks (Figure 3).

Comparison of the chemical shift changes (Δδ) observed for ^{15}N, ^{13}C-labeled ppE7(17–51) in the presence and absence of TAZ2 is shown in Figure S2. Very little change in the ^{13}Cα and ^{13}CO chemical shift is observed, indicating that there is no increase in secondary structure in the E7 peptide in the complex.

A comparison of the chemical shift perturbation (CSP_{E7}) for binding of TAZ2 to ppE7(1–51) and ppE7(17–51) is shown in Figure 4. The CSP values for the CR2 region that is common to the two peptides are quite similar, although those for ppE7(1–51) are consistently higher than those for ppE7(17–51). This observation confirms that both CR1 and CR2 and the intervening residues interact with TAZ2 when ppE7(1–51) binds, and suggests that the CR2 interaction is strengthened when CR1 is present.

Affinity of E7 motifs for TAZ2

To obtain insights into the regions of the E7 sequence that mediate binding to TAZ2 and the relative strength of their interactions, we performed fluorescence anisotropy competition experiments. A complex was formed between TAZ2 and unphosphorylated E7(1–51)(C24A,A50C) labeled with the Alexa594 dye at position 50. The change in fluorescence anisotropy was measured upon addition of competing unlabeled E7 peptides; E7(1–51), E7(17–51), CR1(1–18), and CR2(18–37) (Figure 1A). The peptides were not phosphorylated at S31 and S32 because the primary focus in this experiment was on CR1 and CR2 binding rather than the secondary electrostatic affinity enhancement that occurs upon phosphorylation,[37] and also for the practical reason that competing off the phosphorylated version would have necessitated even higher concentrations of competing peptides. The anisotropy change was fitted to competition models[55] which provided affinity estimates for the individual competing peptides (Figure 5). The results clearly demonstrate that cooperativity between the CR1 and CR2 binding motifs enhances E7 affinity towards TAZ2. The primary CR2 interaction motif has a K_D of 4.6 μM, which is seven-fold stronger than that of the preceding CR1 motif (Table 1). The binding of CR2 is further enhanced 2.5-fold by the presence of residues 42–51 in the peptide E7(17–51), although these residues are unlikely to confer any specificity to the TAZ2 binding. Only when CR1 and CR2 are both present in the peptide is E7 able to reach its characteristic nM binding affinity (K_D 22 nM,[37] compared to 100 nM under the current conditions).

Binding interfaces probed by isotope-filtered NOESY and PRE analysis

Insights into the structure of ppE7(1–51) in complex with TAZ2 were obtained by analyzing the interatomic contacts at the interface. NOE spectra for ^{15}N/^{13}C-TAZ2 in complex with

unlabeled ppE7(1–51) and for $^{15}N/^{13}C$-ppE7(1–51) in complex with unlabeled TAZ2 were acquired with ^{13}C-filters (either filter-edit or edit-filter) to identify either intermolecular (^{12}CH-^{13}CH) or intramolecular (^{13}CH-^{13}CH) NOEs. Representative intermolecular NOEs are listed in Table S1. The NOEs reveal extensive heterogeneity in the complex: I1773 in TAZ2 appears to be close (within ~5 Å) to residues 8, 11, 13, 19, 22, 23, and 25 of ppE7(1–51). The same experiment shows that V1802 is in close proximity to residues 8, 11, 13, 15, 22, 23, 24, and 25. These contacts are not consistent with a single structure. In addition, the leucine methyl and tyrosine aromatic resonances of CR1 and CR2 are heavily overlapped, which not only complicates the structural analysis, but provides further evidence that these residues do not participate in distinct single structures, for which distinctive chemical environments and hence distinctive chemical shifts would be expected. These side chain resonances are only slightly shifted from their chemical shifts in the free peptide spectra, consistent with a high degree of conformational averaging in the TAZ2 complex.

In order to determine whether there was a preferred orientation of the E7 peptide in the complex, we used paramagnetic relaxation enhancement (PRE). Spin-labeled samples of ppE7(1–51) were prepared by first mutating the native cysteine 24 to alanine, and engineering a cysteine in place of the histidine at position 9, within the TAZ2 binding motif in CR1 (Figure 1A). A paramagnetic nitroxide, [2,2,6,6-tetramethylpiperidin-1-yl]oxyl (TEMPO), was coupled via maleimide chemistry to residue 9 (in CR1) of ppE7(1–51) carrying the mutations H9C and C24A. If the unpaired electron in the nitroxide is close to a proton in ^{15}N-TAZ2, the corresponding cross peak in the ^{1}H-^{15}N-HSQC spectrum should broaden to an extent that depends on the interatomic distance,[56] thus locating interactions between TAZ2 and CR1. Figure 6 shows the ^{1}H-^{15}N HSQC spectrum for ^{15}N-TAZ2:ppE7(1–51)(H9C,C24A)-TEMPO in a 1:1 complex. If the spin-labeled E7 took up a preferred orientation within the TAZ2 complex, we would expect broadening of only a subset of peaks within the TAZ2 HSQC spectrum. Instead, as seen in Figure 6 (red), almost every cross peak in the TAZ2 spectrum is broadened by the paramagnetic spin label in the 1:1 complex. The intensity of these resonance signals returns upon the addition of ascorbic acid, which reduces the spin label and removes its relaxation effects (Figure 6, black). Such extensive broadening could not occur unless CR1 interacts with multiple sites on TAZ2. The weak cross peaks remaining in the spectrum of TAZ2 in complex with the spin-labeled peptide belong mostly to residues in the α1-α2 and α3-α4 loops (Figure 6, inset); these sites must be the most distant from the interaction sites of the CR1 motif. The widespread broadening caused by the paramagnetic spin label in the CR1 binding motif shows clearly that the complex is highly disordered and dynamic. Consistent with the results of the filter-edit NOE analysis, the PREs confirm that the CR1 domain of E7(1–51) interacts with multiple regions of the TAZ2 surface, rather than making specific contacts in a localized binding site. Taken together, these results strongly suggest a high degree of disorder in the interactions between the disordered 51-residue N-terminus of E7 and TAZ2.

Contributions of conserved motifs of E7(1–51) to TAZ2 binding

The interaction surface on TAZ2 was identified by following changes in the ^{1}H-^{15}N HSQC spectrum of ^{15}N-labeled TAZ2 (Figure S3) upon addition of CR1, CR2, and E7(1–51) to saturating peptide:TAZ2 ratios. All of the peptides bind to TAZ2 in fast exchange on the

chemical shift timescale, and the titration data were used to determine the NMR-derived K_d values (Table 1) and the saturating peptide concentrations. The fitted NMR titration data for CR1 and CR2 are shown in Figure S4.

The extent to which the cross peak of a given TAZ2 residue moves upon addition of peptide is termed the chemical shift perturbation of TAZ2 (CSP_{TAZ}) and is calculated as $\Delta\delta_{ave} = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{1/2}$. The CSP_{TAZ} as a function of TAZ2 residue number (Figure 7) revealed surprising similarities in the binding of the three peptide constructs to TAZ2.

Figure 7 shows that the overall pattern of TAZ2 chemical shift changes upon binding the CR1 or CR2 peptides is very similar. Some resonances (e.g. L1780 and L1826) experience the same CSP_{TAZ} in the presence of the CR1, CR2, or E7(1–51) peptide constructs while others (e.g. V1819, Y1829) shift to a similar extent in the presence of CR1 and CR2 but are shifted further by binding of E7(1–51). The chemical shift changes are broadly distributed across the α1-α2-α3 and α3-α4 interfaces, and no region of TAZ2 seems affected preferentially by CR1 or CR2. The few differences point to a CR2-dominated change at residues 1800–1802, and a CR1-specific change at L1823. Interestingly, all major CSP_{TAZ} for both CR1 and CR2 are located at the same central binding surface as is contacted by E7(1–51) (Figure 1C), suggesting that CR1 and CR2 can both bind at the same extensive interaction site. This is supported by the observation that addition of CR1 to a CR2:TAZ2 complex does not lead to additional changes outside the binding region. Small shifts are seen in the cross peaks that were affected by CR1 in its binary interaction with TAZ2 (Figure S5). Based on the binary K_Ds of the complexes TAZ2-CR1 (69 ± 7 μM by NMR titration, 31± 2 μM by fluorescence anisotropy competition) and TAZ2-CR2 (3.4 ± 0.5 μM by NMR titration, 4.6 ± 0.5 μM by fluorescence anisotropy competition) (Figure S4, Table 1), we would expect the competition to favor binding of CR2 over CR1. However, the contribution of both CR1 and CR2 together fails to account for all CSP_{TAZ} observed in the E7(1–51) complex. The predominant difference between the binding of CR1/CR2 and that of E7(1–51) is observed for residues 1845–1851 in helix 4 of TAZ2, which undergo larger CSP_{TAZ} upon binding E7(1–51). This could reflect additional interactions when both CR1 and CR2 are present or possibly changes in the length and/or stability of helix 4 of TAZ2 upon binding of the longer peptide.

An attempt was made to stabilize a discrete complex by fusing E7(1–51) to either the N-terminus [$E7(1–51)_{FUS}TAZ2$] or C-terminus [$TAZ2_{FUS}E7(1–51)$] of the TAZ2 domain (Figure 8A), using linkers of various lengths (either no additional residues, a GS linker or a GSGSGS linker). Because these termini are relatively close in space and allow for intramolecular E7 binding within the fusion protein, we wanted to explore if the positioning of E7 could affect the equilibrium between CR1 and CR2 binding. There is a slight preference for the resonances of the resulting complex between the disordered N-terminal 51 residues of E7 and TAZ2 to be affected by whichever domain (CR1 or CR2) is in closest proximity (Figure 8B), but in general we observe that both fusions affect the TAZ2 spectrum very similarly. The cross peaks corresponding to the fused E7 peptide (which is also isotopically labeled when the fusion is expressed for NMR spectroscopy) are also all rather similar to those of the complex of labeled E7(1–51) with unlabeled TAZ2 (Figure 8C).

Structural heterogeneity of E7 in complex with TAZ2

The results described above, including the K_Ds for the CR1 and CR2 complexes shown in Table 1, indicate that TAZ2 has a small preference for binding to CR2 compared to CR1. We therefore proceeded with a structural investigation of the E7(17–51):TAZ2 interaction to elucidate how the CR2 domain contacted the TAZ surface. The free E7(17–51) peptide is disordered in solution, and the small differences in the ^{13}CO and ^{13}Cα chemical shifts between free and bound peptide (Figure S2) indicate that no stable secondary structure is formed by the peptide in the complex. Using a sample containing TAZ2 labeled with ^{13}C in complex with unlabeled ppE7(17–51), we recorded intra- and intermolecular NOESY spectra, in order to obtain distance restraints for structural modeling. Conventionally, such filtered spectra enable unambiguous assignment of contacts between the two binding partners in question. Filtered NOE spectra of the TAZ2 complex of the full peptide ppE7(1–51) showed a number of high-intensity intermolecular NOEs, but the severe overlap of the E7 leucine and tyrosine side chain protons precluded unambiguous assignment to individual residues. For the ppE7(17–51):TAZ2 complex, we expected that the absence of the competing CR1 domain would simplify both the intermolecular contacts and the interpretation of the spectroscopic data. However, as illustrated in Figure 9A and Figure S6, we detected strong NOEs from the overlapped L22/28 methyls and Y23/Y25 aromatic protons to a variety of residues that are widely separated on TAZ2. For example, the two E7 tyrosines are only separated by one residue, yet they contact I1773, T1813, V1802 and I1847, which are located on different surfaces of TAZ2 and are as much as 20 Å apart (Figure 9B). A similar surface is contacted by L22 and/or L28 (Figure 9C). Such heterogeneity of intermolecular contacts is inconsistent with the presence of a single binding mode for ppE7(17–51) with TAZ2 but indicates averaging between multiple binding poses over an extensive hydrophobic surface of TAZ2.

It is clear from all of the results shown in Figures 2, 3, 6 and 9 that no single E7 structure will be able to satisfy all of the NOE constraints. To obtain insights into potential binding modes, we used the Haddock molecular docking web server[54] and the TAZ2 structure from the E1A:TAZ2 complex (PDB: 2KJE[41]) as input. For E7 we used a SAXS-constrained molecular ensemble of 20 models generated by the Allosmod-FOXS server[52,53] using experimental SAXS data collected on E7(1–51) (Figure S7). We specified a range of distance constraints based on intermolecular NOEs observed for the CR2 region of the E7(1–51):TAZ2 complex (examples in Figure S6). The E7 input models consisted only of residues 17–40 because no interactions with TAZ2 were observed for residues 41–51. Convergence of the Haddock models was problematic due to the contradictory nature of much of the NOE data; no single structure could satisfy all of the NOE constraints. Convergence could be achieved by using subsets of the NOEs in separate calculations, giving rise to a wide variety of docked structures, with the peptide in multiple conformations and poses. However, the heterogeneity of the interaction precludes a more quantitative evaluation of the structural ensemble.

DISCUSSION

The flexibility of IDPs and IDRs is an important factor in their universal employment in cellular signaling, cellular regulation and molecular recognition.[6–9] IDP interactions with folded or unfolded partners frequently include binding-induced folding of the IDP.[57,58] However, an increasing number of protein complexes involving IDPs have been found to involve structural heterogeneity and flexibility; these complexes (often termed fuzzy complexes[6,59]) cannot be described by a single conformational state, and such fuzzy complexes may have exceptionally high affinities.[60] Viral proteins like E7 must interact with a myriad of host-cell binding partners in order to achieve cellular hijacking with the minimal genetic information of the virus. The ability of E7 to interact with a wide array of cellular partners originates primarily from its intrinsically disordered N-terminal domain.[61]

Extensive analysis of the disordered N-terminal region of E7, both in isolation and in the context of the full-length protein,[37,62,63] indicates that it is conformationally heterogeneous and samples both extended and compact states. Interactions between the dimeric full-length E7 and the TAZ2 domain of CBP are in intermediate exchange on the NMR timescale and result in severe resonance broadening,[37] precluding structural analysis. We therefore focused on the intrinsically disordered N-terminal region of E7, which binds TAZ2 with nanomolar K_D via its CR1 and CR2 motifs,[37] with the goal of characterization of the complex between the 51-residue peptide encompassing the intrinsically disordered N-terminal region of E7 (Figure 1A) and the TAZ2 domain of CBP/p300. While there are examples of high-resolution structures for several IDPs in complex with TAZ2,[41,64–69] the TAZ2-E7 system has presented a more challenging case. NMR experiments of the complex of ppE7(1–51) and TAZ2 revealed that the peptide binds in fast exchange on the chemical shift timescale and remains highly dynamic in the complex, despite its nanomolar binding affinity for TAZ2.[37] Unlike E1A from adenovirus, which forms helical segments upon binding to TAZ2,[41] there is minimal difference in secondary structure between the bound and unbound states of E7(1–51).

Synergy between CR1 and CR2 in complex with TAZ2

The significant chemical shift differences seen in the ^{1}H-^{15}N HSQC spectra between free and bound E7(1–51) and the observation of intermolecular NOEs between TAZ2 and residues in both the CR1 and CR2 motifs clearly demonstrate that both domains are in contact with TAZ2 in the complex. Many of the intermolecular NOEs are mutually incompatible; this implies that residues within the helix α_1-α_2-α_3 interface of TAZ2, such as I1773, M1799 and V1802, contact both CR1 and CR2 of E7 in different binding modes. Introduction of a spin label in CR1 resulted in broadening of virtually all TAZ2 cross peaks, also consistent with multiple binding modes. NMR titrations with the isolated CR1 and CR2 peptides show that these motifs bind promiscuously at multiple sites on TAZ2. This is perhaps unsurprising in view of the strong similarities in their amino acid sequences, with the residues ^{21}DLYCYEQL^{28}, which contains the LXCXE motif in CR2 almost mirrored in the reversed CR1 sequence, ^{15}LDLMYEHL^{8}.

In the context of E7(1–51), the CR1 and CR2 motifs interact synergistically with TAZ2. Once one motif binds, say CR2 which has higher intrinsic affinity as an isolated motif, then

the high effective local concentration of CR1 will facilitate its interaction with a neighboring site on TAZ2. Chemical shift perturbations (Figure 4) show that residues 18–21, which form the linker between CR1 and CR2, contact TAZ2 in the E7(1–51) complex whereas they do not contribute to binding by the isolated CR2 peptide or E7(17–51). Simultaneous binding of the CR1 and CR2 motifs, albeit through dynamically disordered interactions, would explain the observed increase in the overall binding affinity of E7(1–51) relative to the isolated CR1 and CR2 peptides (Table 1). We may compare the behavior of E7(1–51) with that of the p53 transactivation domain. The p53 TAD contains interaction sites, termed AD1 and AD2, which both bind preferentially to the TAZ2 α1-α2-α3 site as isolated peptides. Both peptides have a secondary binding site in the α3-α4 groove. However, when both are present in the full-length TAD, AD2 binds in the preferred α1-α2-α3 site while AD1 binds to the α3-α4 groove.[55] For p53, there is a large difference in the K_d for binding to the primary α1-α2-α3 site: 32 nM for AD2 vs 24 μM for AD1, consistent with the binding preference for AD2 when both are present. For E7, the K_ds for the isolated CR1 and CR2 domains are more similar than for p53 and both forward and reverse binding poses, with both CR1 and CR2 occupying all possible TAZ2 binding sites, will be populated. This is confirmed by the intermolecular NOEs observed for the complex with E7(1–51); CR1 and CR2 residues both exhibit NOEs to a cluster of residues in the α1-α2-α3 site (Table S1) showing that each of the motifs contacts this region of TAZ2 in different structures in the conformational ensemble.

Conformational heterogeneity as an effective binding strategy

The interaction between the disordered region of E7 and the TAZ2 domain of CBP operates through a multiplicity of intermolecular contacts, centered around key hydrophobic residues. Although the affinity between the two molecules is high, binding is in fast exchange on the NMR timescale (Figure 2). This may be rationalized by two general circumstances: firstly, there is a general overall electrostatic attraction at pH 7 between the TAZ2 domain (net charge +12.6) and E7(1–51) (net charge −13.4). Secondly, binding of both CR1 and CR2 to the extensive interaction surface on TAZ2 will enhance the overall binding affinity of E7(1–51) relative to the isolated domains. The complex is heterogeneous and highly dynamic: either CR1 or CR2 can occupy the α1-α2-α3 site, with the other motif interacting elsewhere on the surface of TAZ2 and contributing to the binding free energy. Despite the high overall affinity, the interactions of the individual CR1 and CR2 motifs with TAZ2 are relatively weak and they exchange rapidly between bound and free states. Mechanistically, E7 association and dissociation is expected to occur via an intermediate state in which only the CR1 or CR2 motif is bound to TAZ2 and since binding of each is in fast exchange on the chemical shift time scale, the overall exchange process for E7 binding is also fast. High affinity is accomplished through multivalent interactions, with binding of the individual motifs being relatively weak and with no requirement for specificity. The lack of specific interactions, and the promiscuity of the binding (many binding sites, many binding poses) acts to strengthen the overall affinity of the disordered ligand for its partner. The validity of the model is attested by the observation that there is very little chemical shift dispersion in the ^{1}H-^{13}C HSQC spectrum of E7 in complex with TAZ2 (Figure 3): the dynamic heterogeneity of the complex, and the fast exchange between free and bound states

averages the chemical shifts of side chains such as L22, L28, Y23 and Y25, which would be well-dispersed in a complex with a defined structure.

Under the experimental conditions, unphosphorylated E7 has a net negative charge (−13.4), while TAZ2 has a net positive charge (+12.6). The overall electrostatic attraction between the two molecules keeps them in close proximity, allowing alternative regions of the peptide to bind promiscuously throughout the TAZ2 binding surface. The effect of phosphorylation of the two serine residues in E7(1–51) provides corroboration for this picture. Chemical shift differences between the free and bound E7 peptide[37] (Figure 4) indicate that E7 binds to TAZ2 in the region between residues Y11 and D30. The phosphorylated serines do not appear to make direct contact with TAZ2, yet phosphorylation increases the affinity of E7(1–51) for TAZ2,[37] thus confirming the role of the overall charge difference in the mechanism of E7-TAZ2 binding.

Significance of fuzzy complexes for viruses

The employment of intrinsic disorder and formation of fuzzy complexes emerges as a suitable strategy for E7 and other similar viral proteins that target a myriad of cellular processes to reprogram the host cell to favor viral replication. Given their limited genomes, viruses such as HPV need to utilize highly versatile proteins to maximize their functional impact. In high-risk HPV, one of the primary roles of E7 is to interfere with transcription and dysregulate the cell cycle,[16,19,23] which it does by interacting with approximately thirty different host cell proteins, including the retinoblastoma protein pRb and CBP/p300. Almost half of the known partners bind primarily to the LXCXE motif.[61] Very few of these interactions have been structurally characterized.[34,62] The E7-pRb interaction results in a well-defined complex,[34] whereas we show that the interaction of the same peptide with TAZ2 is highly dynamic and conformationally heterogeneous. Interestingly, the affinities of E7(1–51) for pRb and for TAZ2 are comparable, resulting in the formation of a mixture of complexes between the disordered 51-residue N-terminus of E7 and either pRb or TAZ2 when all three components are mixed.[37] The formation of dynamic and heterogeneous complexes is likely advantageous for the virus. By utilizing promiscuous hydrophobic interactions and perhaps charge complementarity, E7 can dysregulate host cell processes through high affinity, fuzzy interactions with numerous cellular proteins. Specific interactions, mediated through defined structure, are unnecessary. In this way, the virus can utilize a single protein, encoded by its limited genome, to bind promiscuously to numerous host cell proteins, without a requirement for evolution of highly specific interaction motifs that mimic cellular recognition processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Gerard Kroon for expert assistance with NMR experiments, Rebecca Berlow for assistance with data analysis, Euvel Manlapaz for technical assistance, Tsutomu Matsui for assistance with SAXS data collection at the SLAC beam line, and Logan Morin, Bailey Holmes and other members of the Jansma lab for valuable discussions. This work was supported by grants CA096865 (PEW) and GM131693 (HJD) from the National Institutes of Health

and by the Skaggs Institute for Chemical Biology (PEW). MWR was supported by the Carlsberg Foundation's Internationalization Fellowship. AJ was supported by Point Loma Nazarene University Research Associates.

REFERENCES

[1]. Wright PE, and Dyson HJ (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm, J. Mol. Biol. 293, 321–331. [PubMed: 10550212]

[2]. Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, and Dunker AK (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins, J. Mol. Biol 323, 573–584. [PubMed: 12381310]

[3]. Tompa P (2002) Intrinsically unstructured proteins, Trends Biochem. Sci. 27, 527–533. [PubMed: 12368089]

[4]. Dyson HJ, and Wright PE (2005) Intrinsically unstructured proteins and their functions, Nature Rev. Mol. Cell Biol 6, 197–208. [PubMed: 15738986]

[5]. Wright PE, and Dyson HJ (2015) Intrinsically disordered proteins in cellular signalling and regulation, Nature Rev. Mol. Cell Biol 16, 18–29. [PubMed: 25531225]

[6]. Tompa P, and Fuxreiter M (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions, Trends Biochem. Sci. 33, 2–8. [PubMed: 18054235]

[7]. Uversky VN (2018) Intrinsic disorder, protein–protein interactions, and disease, Adv. Prot. Chem. Struct. Biol. 110, 85–121.

[8]. Fuxreiter M (2020) Classifying the binding modes of disordered proteins, Int. J. Mol. Sci. 21, 8615.

[9]. Miskei M, Horvath A, Vendruscolo M, and Fuxreiter M (2020) Sequence-Based Prediction of Fuzzy Protein Interactions, J. Mol. Biol. 432, 2289–2303. [PubMed: 32112804]

[10]. Garcia-Alai MM, Alonso LG, and de Prat-Gay G (2007) The N-Terminal module of HPV16 E7 is an intrinsically disordered domain that confers conformational and recognition plasticity to the oncoprotein, Biochemistry 46, 10405–10412. [PubMed: 17715947]

[11]. Xue B, Williams RW, Oldfield CJ, Goh GK, Dunker AK, and Uversky VN (2010) Viral disorder or disordered viruses: do viral proteins possess unique features? Prot. Pept. Lett 17, 932–951.

[12]. Borkosky SS, Camporeale G, Chemes LB, Risso M, Noval MG, Sánchez IE, Alonso LG, and de Prat Gay G (2017) Hidden structural codes in protein intrinsic disorder, Biochemistry 56, 5560–5569. [PubMed: 28952717]

[13]. Tamarozzi E, and Giuliatti S (2018) Understanding the role of intrinsic disorder of viral proteins in the oncogenicity of different types of HPV, Int. J. Mol. Sci. 19, 198.

[14]. Van Doorslaer K, Li Z, Xirasagar S, Maes P, Kaminsky D, Liou D, Sun Q, Kaur R, Huyen Y, and McBride AA (2017) The Papillomavirus Episteme: a major update to the papillomavirus sequence database, Nucl. Acids Res. 45, D499–D506. [PubMed: 28053164]

[15]. de Villiers E-M, Fauquet C, Broker TR, Bernard H-U, and zur Hausen H (2004) Classification of papillomaviruses, Virology 324, 17–27. [PubMed: 15183049]

[16]. McBride AA (2017) Oncogenic human papillomaviruses, Phil. Trans. Roy. Soc. B: Mathematics, Physics and Engineering Science 372.

[17]. McBride AA (2017) Playing with fire: consequences of human papillomavirus DNA replication adjacent to genetically unstable regions of host chromatin, Curr. Opin. Virol 26, 63–68. [PubMed: 28779692]

[18]. Stewart BW, and Wild CP (2014) World Cancer Report, Vol. 3, International Agency for Research on Cancer, Lyon, France.

[19]. Ganguly N, and Parihar SP (2009) Human papillomavirus E6 and E7 oncoproteins as risk factors for tumorigenesis, J. Biosci. 34, 113–123. [PubMed: 19430123]

[20]. Helt A-M, and Galloway DA (2003) Mechanisms by which DNA tumor virus oncoproteins target the Rb family of pocket proteins, Carcinogenesis 24, 159–169. [PubMed: 12584163]

[21]. Münger K, Phelps WC, Bubb V, Howley PM, and Schlegel R (1989) The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes, J. Virol. 63, 4417–4421. [PubMed: 2476573]

[22]. Levine AJ (2009) The common mechanisms of transformation by the small DNA tumor viruses: The inactivation of tumor suppressor gene products: p53, Virology 384, 285–293. [PubMed: 19081592]

[23]. McLaughlin-Drubin ME, and Münger K (2009) The human papillomavirus E7 oncoprotein, Virology 384, 335–344. [PubMed: 19007963]

[24]. Ohlenschläger O, Seiboth T, Zengerling H, Briese L, Marchanka A, Ramachandran R, Baum M, Korbas M, Meyer-Klaucke W, Dürst M, and Görlach M (2006) Solution structure of the partially folded high-risk human papilloma virus 45 oncoprotein E7, Oncogene 25, 5953–5959. [PubMed: 16636661]

[25]. Calçada EO, Felli IC, Hosek T, and Pierattelli R (2013) The heterogeneous structural behavior of E7 from HPV16 revealed by NMR spectroscopy, Chembiochem 14, 1876–1882. [PubMed: 23940009]

[26]. Noval MG, Gallo M, Perrone S, Salvay AG, Chemes LB, and De Prat-Gay G (2013) Conformational dissection of a viral intrinsically disordered domain involved in cellular transformation, PLoS ONE 8, e72760. [PubMed: 24086265]

[27]. Clemens KE, Brent R, Gyuris J, and Münger K (1995) Dimerization of the human papillomavirus E7 oncoprotein in vivo, Virology 214, 289–293. [PubMed: 8525630]

[28]. Liu X, Clements A, Zhao K, and Marmorstein R (2006) Structure of the human papillomavirus E7 oncoprotein and its mechanism for inactivation of the retinoblastoma tumor suppressor, J. Biol. Chem. 281, 578–586. [PubMed: 16249186]

[29]. Todorovic B, Massimi P, Hung K, Shaw GS, Banks L, and Mymryk JS (2011) Systematic analysis of the amino acid residues of human papillomavirus type 16 E7 conserved region 3 involved in dimerization and transformation, J. Virol. 85, 10048–10057. [PubMed: 21775462]

[30]. Mirabello L, Yeager M, Yu K, Clifford GM, Xiao Y, Zhu B, Cullen M, Boland JF, Wentzensen N, Nelson CW, Raine-Bennett T, Chen Z, Bass S, Song L, Yang Q, Steinberg M, Burdett L, Dean M, Roberson D, Mitchell J, Lorey T, Franceschi S, Castle PE, Walker J, Zuna R, Kreimer AR, Beachler DC, Hildesheim A, Gonzalez P, Porras C, Burk RD, and Schiffman M (2017) HPV16 E7 genetic conservation is critical to carcinogenesis, Cell 170, 1164–1174.e1166. [PubMed: 28886384]

[31]. Chemes LB, Glavina J, Alonso LG, Marino-Buslje C, De Prat-Gay G, and Sanchez IE (2012) Sequence evolution of the intrinsically disordered and globular domains of a model viral oncoprotein, PLoS ONE 7, e47661. [PubMed: 23118886]

[32]. Heck DV, Yee CL, Howley PM, and Münger K (1992) Efficiency of binding the retinoblastoma protein correlates with the transforming capacity of the E7 oncoproteins of the human papillomaviruses, Proc. Natl Acad. Sci. USA 89, 4442–4446. [PubMed: 1316608]

[33]. Jones DL, Thompson DA, and Münger K (1997) Destabilization of the RB tumor suppressor protein and stabilization of p53 contribute to HPV type 16 E7-induced apoptosis, Virology 239, 97–107. [PubMed: 9426450]

[34]. Lee JO, Russo AA, and Pavletich NP (1998) Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7, Nature 391, 859–865. [PubMed: 9495340]

[35]. Suarez I, and Trave G (2018) Structural insights in multifunctional papillomavirus oncoproteins, Viruses 10, 37.

[36]. De Guzman RN, Liu HY, Martinez-Yamout M, Dyson HJ, and Wright PE (2000) Solution structure of the TAZ2 (CH3) domain of the transcriptional adaptor protein CBP, J. Mol. Biol. 303, 243–253. [PubMed: 11023789]

[37]. Jansma AL, Martinez-Yamout MA, Liao R, Sun P, Dyson HJ, and Wright PE (2014) The high-risk HPV16 E7 oncoprotein mediates interaction between the transcriptional coactivator CBP and the retinoblastoma protein pRb, J. Mol. Biol. 426, 4030–4048. [PubMed: 25451029]

[38]. Goodman RH, and Smolik S (2000) CBP/p300 in cell growth, transformation, and development, Genes Devel. 14, 1553–1577. [PubMed: 10887150]

[39]. Chan HM, and La Thangue NB (2001) p300/CBP proteins: HATs for transcriptional bridges and scaffolds, J. Cell Sci. 114, 2363–2373. [PubMed: 11559745]

[40]. Dyson HJ, and Wright PE (2016) Role of intrinsic protein disorder in the function and interactions of the transcriptional coactivators CREB-binding protein (CBP) and p300, J. Biol. Chem. 291, 6714–6722. [PubMed: 26851278]

[41]. Ferreon JC, Martinez-Yamout MA, Dyson HJ, and Wright PE (2009) Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein, Proc. Natl Acad. Sci USA 106, 13260–13265. [PubMed: 19651603]

[42]. De Guzman RN, Wojciak JM, Martinez-Yamout MA, Dyson HJ, and Wright PE (2005) CBP/p300 TAZ1 domain forms a structured scaffold for ligand binding, Biochemistry 44, 490–497. [PubMed: 15641773]

[43]. Lundblad JR, Laurance M, and Goodman RH (1996) Fluorescence polarization analysis of protein-DNA and protein-protein interactions, Mol. Endocrinol. 10, 607–612. [PubMed: 8776720]

[44]. Lee CW, Ferreon JC, Ferreon AC, Arai M, and Wright PE (2010) Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation, Proc. Natl Acad. Sci. USA 107, 19290–19295. [PubMed: 20962272]

[45]. Lee W, Revington MJ, Arrowsmith C, and Kay LE (1994) A pulsed field gradient isotope-filtered 3D ^{13}C HMQC-NOESY experiment for extracting intermolecular NOE contacts in molecular complexes, FEBS Lett. 350, 87–90. [PubMed: 8062930]

[46]. Shen Y, and Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, J. Biomol. NMR 56, 227–241. [PubMed: 23728592]

[47]. Deshmukh L, Tugarinov V, Louis JM, and Clore GM (2017) Binding kinetics and substrate selectivity in HIV-1 protease−Gag interactions probed at atomic resolution by chemical exchange NMR, Proc. Natl Acad. Sci. USA 114, E9855–E9862. [PubMed: 29087351]

[48]. Lee CW, Arai M, Martinez-Yamout MA, Dyson HJ, and Wright PE (2009) Mapping the interactions of the p53 transactivation domain with the KIX domain of CBP, Biochemistry 48, 2115–2124. [PubMed: 19220000]

[49]. McPhillips TM, McPhillips SE, Chiu HJ, Cohen AE, Deacon AM, Ellis PJ, Garman E, Gonzalez A, Sauter NK, Phizackerley RP, Soltis SM, and Kuhn P (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines, J. Synch. Rad. 9, 401–406.

[50]. Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM, and Svergun DI (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions, J. Appl. Crystallog. 50, 1212–1225.

[51]. Tria G, Mertens HD, Kachala M, and Svergun DI (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, International Union of Crystallography Journal 2, 207–217.

[52]. Weinkam P, Pons J, and Sali A (2012) Structure-based model of allostery predicts coupling between distant sites, Proc. Natl Acad. Sci. USA 109, 4875–4880. [PubMed: 22403063]

[53]. Schneidman-Duhovny D, Hammel M, and Sali A (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles, Nucleic Acids Res. 38, W540–544. [PubMed: 20507903]

[54]. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, and Bonvin AMJJ (2016) The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes, J. Mol. Biol. 428, 720–725. [PubMed: 26410586]

[55]. Arai M, Ferreon JC, and Wright PE (2012) Quantitative analysis of multisite protein-ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP, J. Am. Chem. Soc. 134, 3792–3803. [PubMed: 22280219]

[56]. Huang S, Umemoto R, Tamura Y, Kofuku Y, Uyeda TQ, Nishida N, and Shimada I (2016) Utilization of paramagnetic relaxation enhancements for structural analysis of actin-binding proteins in complex with actin, Sci. Rep. 6, 33690. [PubMed: 27654858]

[57]. Wright PE, and Dyson HJ (2009) Linking folding and binding, Curr. Opin. Struct. Biol. 19, 31–38. [PubMed: 19157855]

[58]. Tompa P, Schad E, Tantos A, and Kalmar L (2015) Intrinsically disordered proteins: emerging interaction specialists, Curr. Opin. Struct. Biol. 35, 49–59. [PubMed: 26402567]

[59]. Fuxreiter M, and Tompa P (2009) Fuzzy interactome: the limitations of models in molecular biology, Trends Biochem. Sci. 34, 3–3.

[60]. Borgia A, Borgia MB, Bugge K, Kissling VM, Heidarsson PO, Fernandes CB, Sottini A, Soranno A, Buholzer KJ, Nettels D, Kragelund BB, Best RB, and Schuler B (2018) Extreme disorder in an ultrahigh-affinity protein complex, Nature 555, 61–66. [PubMed: 29466338]

[61]. Songock WK, Kim SM, and Bodily JM (2017) The human papillomavirus E7 oncoprotein as a regulator of transcription, Virus Res. 231, 56–75. [PubMed: 27818212]

[62]. Chemes LB, Sanchez IE, Smal C, and De Prat-Gay G (2010) Targeting mechanism of the retinoblastoma tumor suppressor by a prototypical viral oncoprotein, FEBS J. 277, 973–988. [PubMed: 20088881]

[63]. Kukic P, Lo Piccolo GM, Nogueira MO, Svergun DI, Vendruscolo M, Felli IC, and Pierattelli R (2019) The free energy landscape of the oncogene protein E7 of human papillomavirus type 16 reveals a complex interplay between ordered and disordered regions, Sci. Rep. 9, 5822. [PubMed: 30967564]

[64]. Feng H, Jenkins LMM, Durell SR, Hayashi R, Mazur SJ, Cherry S, Tropea JE, 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. [PubMed: 19217391]

[65]. Wojciak JM, Martinez-Yamout MA, Dyson HJ, and Wright PE (2009) Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains, EMBO J. 28, 948–958. [PubMed: 19214187]

[66]. Bhaumik P, Davis J, Tropea JE, Cherry S, Johnson PF, and Miller M (2014) Structural insights into interactions of C/EBP transcriptional activators with the Taz2 domain of p300, Acta Cryst. D 70, 1914–1921. [PubMed: 25004968]

[67]. Miller Jenkins LM, Feng H, Durell SR, Tagad HD, Mazur SJ, Tropea JE, 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. [PubMed: 25753752]

[68]. 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.e1094. [PubMed: 30099987]

[69]. Krois AS, Ferreon JC, Martinez-Yamout MA, Dyson HJ, and Wright PE (2016) Recognition of the disordered p53 transactivation domain by the transcriptional adapter zinc finger domains of CREB-binding protein, Proc. Natl Acad. Sci. USA 113, E1853–E1862. [PubMed: 26976603]

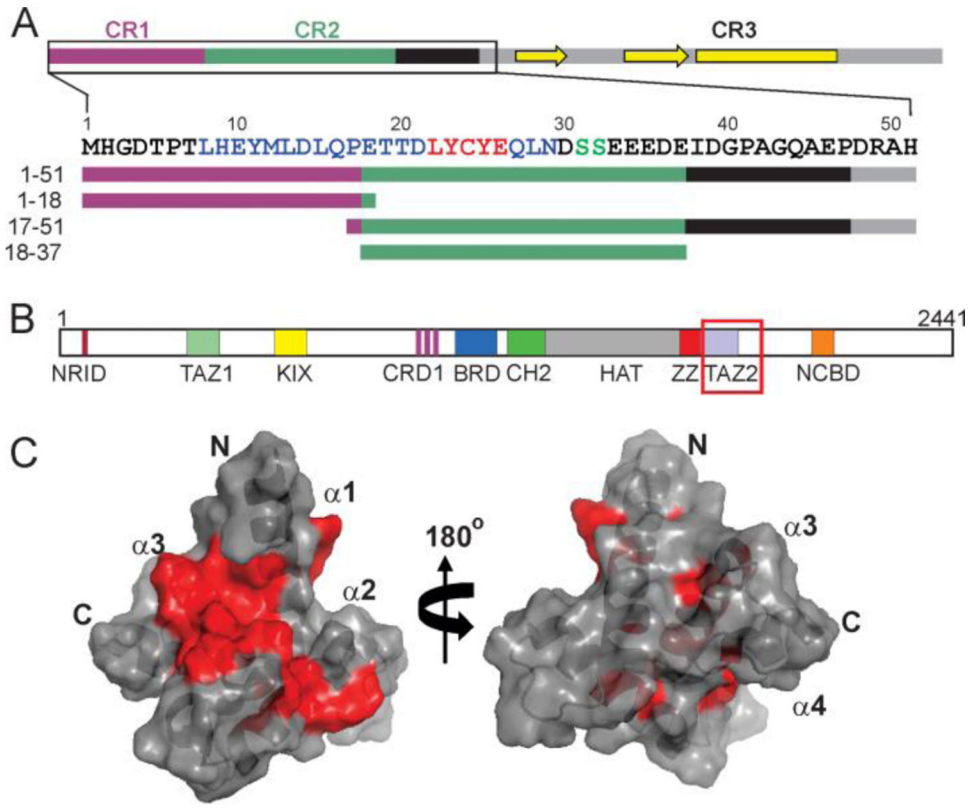


Figure 1.

E7 and CBP TAZ2. A. Schematic diagram of full-length E7 from HPV16 and the amino acid sequence of HPV16(1–51). Conserved regions CR1 (purple), CR2 (green) and CR3 (gray) are shown, with secondary structure elements from the solution structure of the HPV1a CR3 dimer,[24] indicating β-strands (yellow arrows) and α-helix (yellow rectangle). Residues 1–51 are outlined in the black box and expanded to show the amino acid sequence of HPV16(1–51). The TAZ2 binding region is shown in blue and includes the LXCXE motif in red. The sites of serine phosphorylation in CR2 are shown in green. Constructs and their nomenclature designed for the present work are outlined below as colored rectangles. B. domain structure of CBP/p300 showing the TAZ2 domain highlighted in the red box. C. Structure of the TAZ2 domain[36] indicating the binding interface of the disordered region of E7.[37]

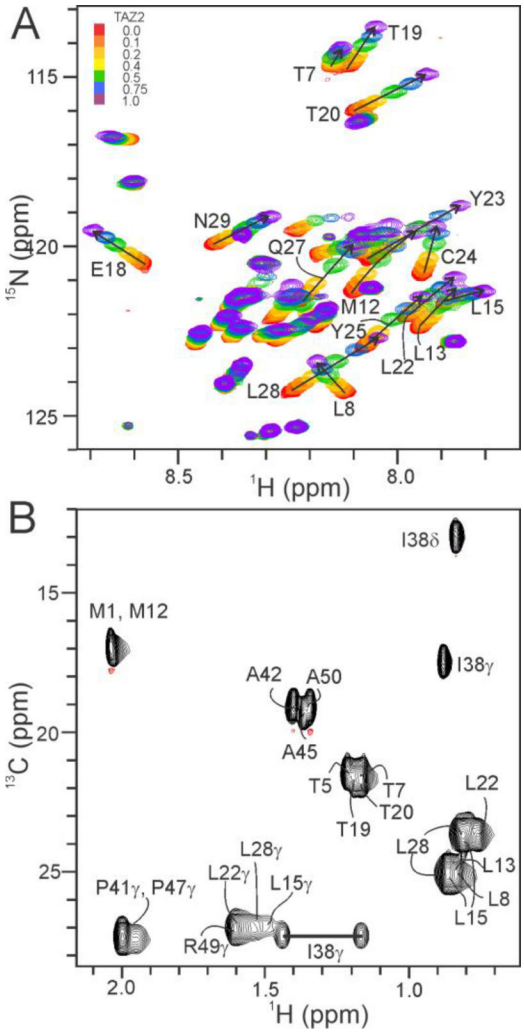


Figure 2.

NMR analysis of ppE7(1–51) in complex with TAZ2. A. ^{1}H-^{15}N-HSQC titration of uniformly labeled ^{15}N-ppE7(1–51) (red) with increasing amounts of TAZ2 up to 1:1 molar ratio (purple). B. Portion of the ^{1}H-^{13}C HSQC spectrum of ^{13}C-labeled ppE7(1–51) in complex with unlabeled TAZ2.

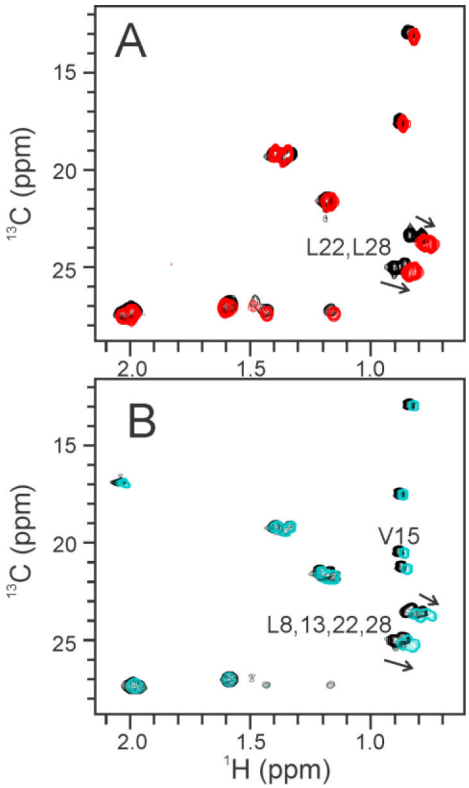


Figure 3.

NMR analysis of ppE7(17–51) and mutant ppE7(1–51) in complex with TAZ2. A. Overlay of the methyl region of the ^{1}H-^{13}C HSQC spectrum of ^{13}C-labeled ppE7(17–51) free (black) and in complex with unlabeled TAZ2 (red). B. Overlay of the methyl region of the ^{1}H-^{13}C HSQC spectrum of ^{13}C-labeled ppE7(1–51)(L15V) free (black) and in complex with unlabeled TAZ2 (cyan).

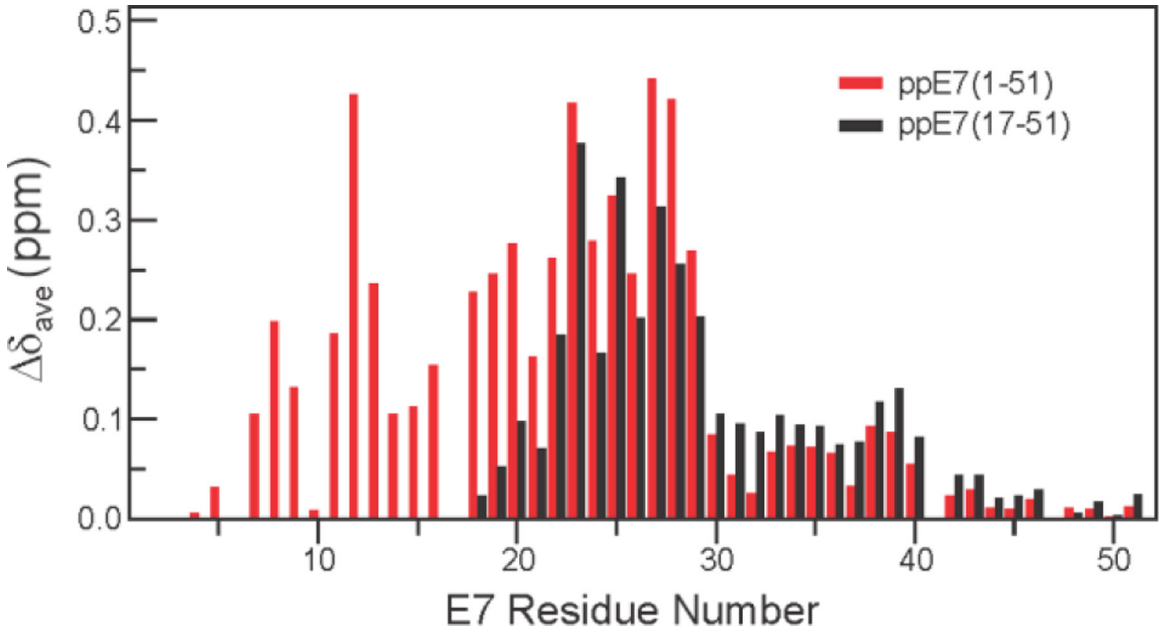


Figure 4.

Plot of weighted average ^{1}H, ^{15}N chemical shift differences ($\Delta\delta_{ave}$) between E7 free and bound to TAZ2. red: ppE7(1–51), (data from ref.[37]); black: ppE7(17–51). Weighted average shifts were calculated using the formula $\Delta\delta_{ave} = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{1/2}$. Note that different weighting was used in.[37]

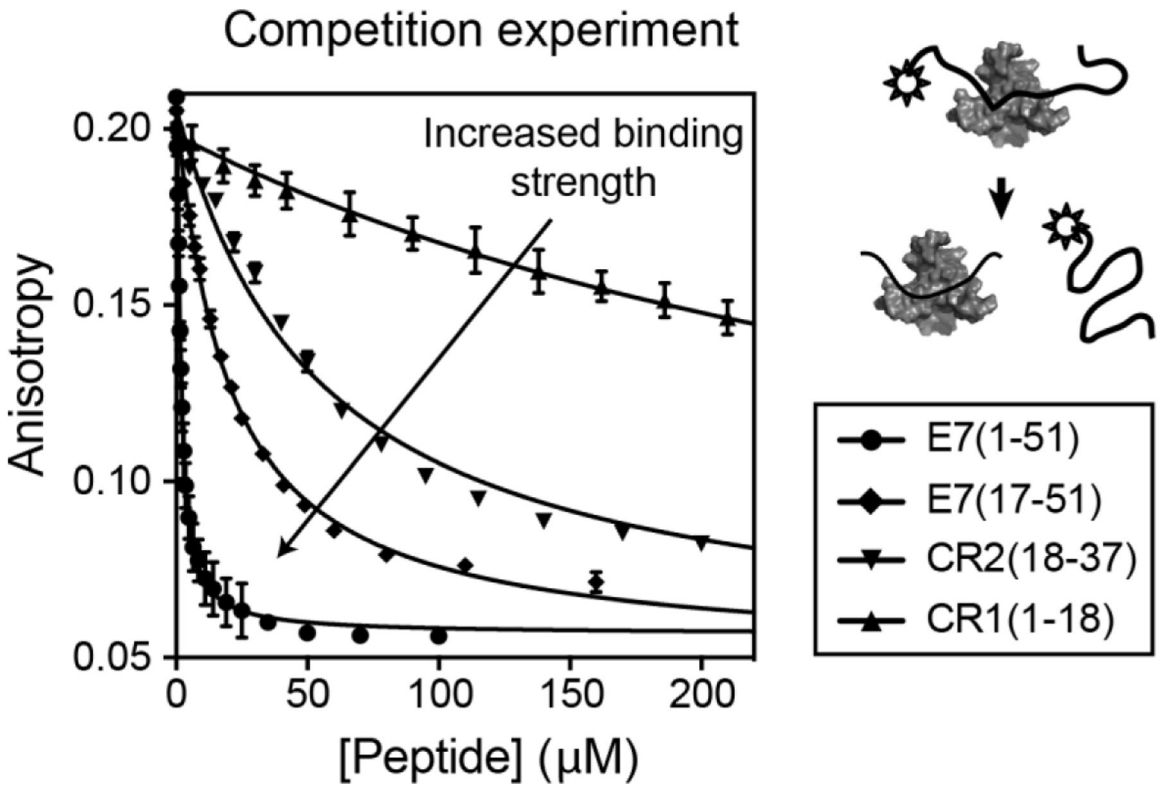


Figure 5.

Fluorescence anisotropy analysis of E7 sequence element binding to TAZ2. The affinity of the E7 peptide constructs towards TAZ2 was calculated from the ability of each construct to compete with Alexa594-labeled E7(1–51) bound to TAZ2. The initial anisotropy value represents the bound Alexa594-E7(1–51) and the decrease in anisotropy as a function of titrant represents an increase in free Alexa594-E7(1–51) caused by a gradual displacement from TAZ2. In this setup, the starting complex Alexa594-E7(1–51):TAZ2 was measured to have a K_D of 79 nM by direct titration. Each competition titration series was performed in triplicate; error bars represent the standard deviation between the three data sets. Curve fitting was performed according to previously published methods.[44,55]

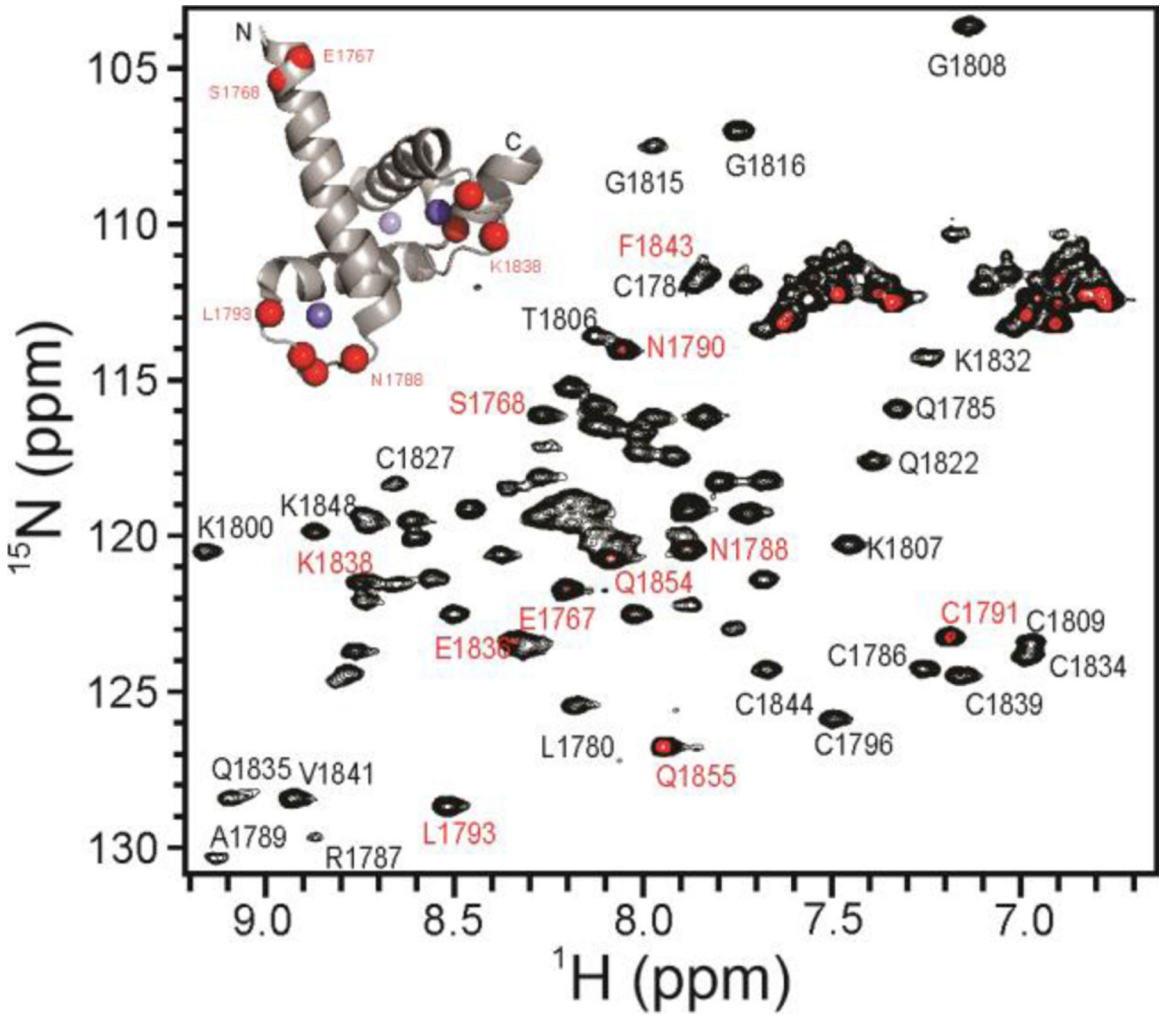


Figure 6.

Paramagnetic relaxation enhancement analysis. ^{1}H-^{15}N HSQC of a 1:1 complex of ^{15}N-TAZ2:ppE7(1–51)(H9C,C24A)-TEMPO (red) and with ascorbic acid added to reduce the nitroxide spin label (black). Cross peaks that are observed in the presence of the active spin label are labeled in red. (Inset) TAZ2 structure showing as red spheres the position of the amide nitrogen of the residues labeled in red in the figure. Zinc atoms are shown as blue spheres.

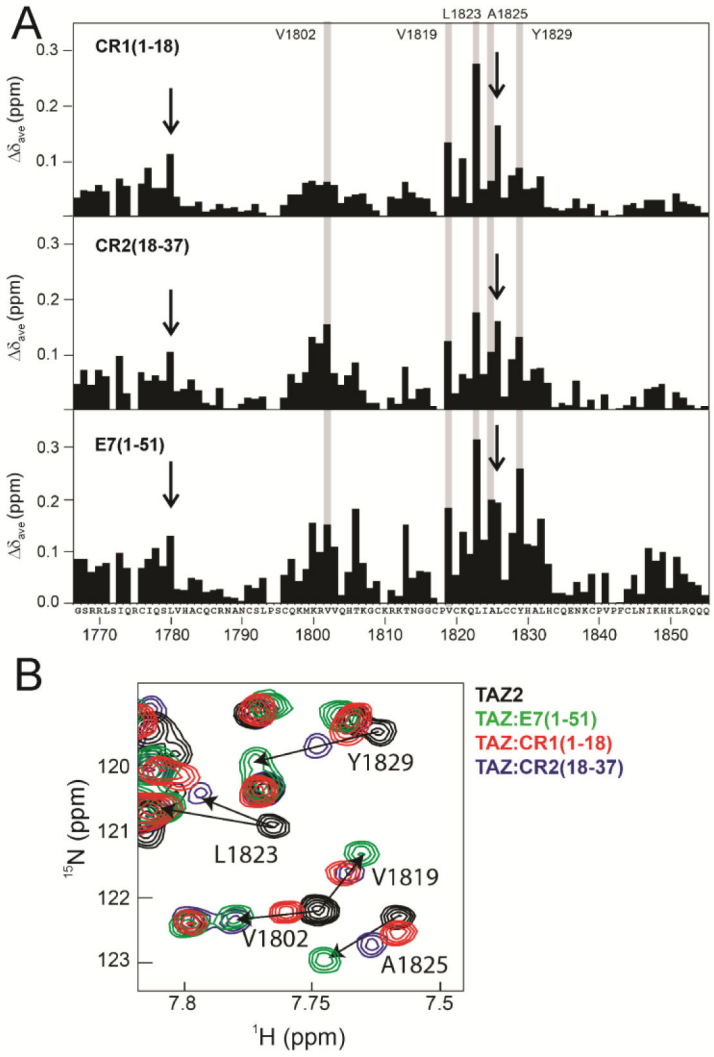


Figure 7.

The isolated CR1 and CR2 domains of E7 share contact interfaces on TAZ2. (A) Chemical shift perturbations CSP_{TAZ} $\Delta\delta_{ave} = [(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2]^{1/2}$ calculated from ^{1}H-^{15}N TAZ2 HSQC spectra recorded in the absence and presence of saturating amounts of E7(1–51), CR1, and CR2 peptides (ratios of 1:1, 1:4.5 and 1:2, respectively). The three E7 peptides display clear similarities in their interaction sites on TAZ2, but differences are observed in the amplitude of the CSP_{TAZ}. Selected residues are highlighted in grey to illustrate differential effects for CR1 and CR2 compared to E7(1–51). Vertical arrows show the positions of residues 1780 and 1826, which are highlighted in the text. (B) Region of the ^{1}H-^{15}N HSQC spectrum of TAZ2 showing perturbations due to binding of CR1, CR2, and E7(1–51).

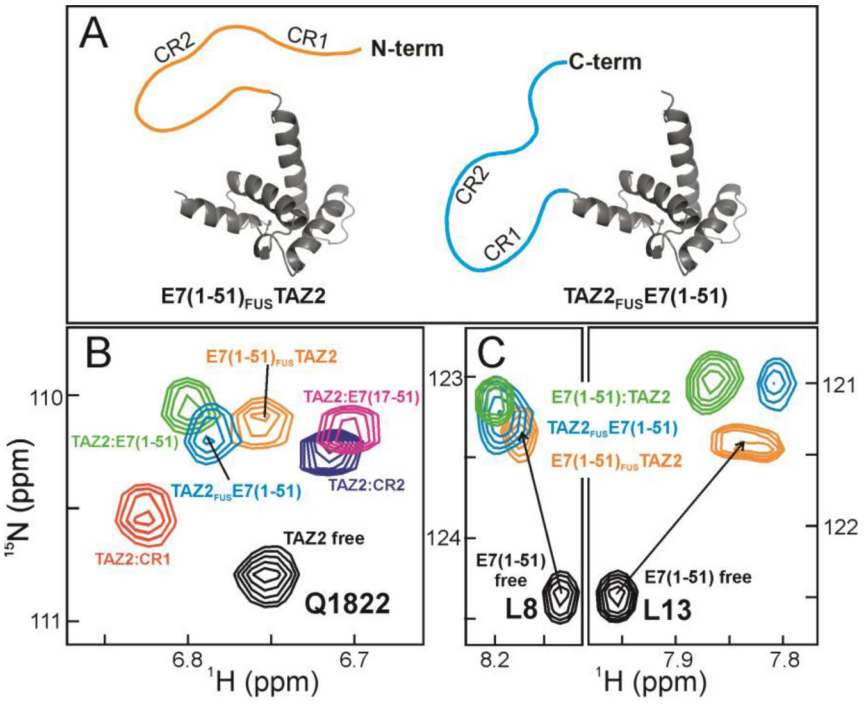


Figure 8.

TAZ2-E7 fusions. (A) Schematic diagram showing the fusion proteins generated to study CR1/CR2 binding to TAZ2. (B) Portion of the ^{1}H-^{15}N TAZ2 HSQC spectrum showing the superposition of the cross peak of the Q1822 side chain for the fusions and free TAZ2 and other TAZ2:peptide complexes. (C) Portion of the ^{1}H-^{15}N E7(1–51) HSQC spectrum showing the superpositions of the cross peaks of the L8 and L13 backbone amides for the fusions and free E7(1–51) and its TAZ2 complex. Arrows show that in E7(1–51)$_{FUS}$TAZ2, the L8 and L13 resonances experience smaller perturbations from the free peptide positions, since the C-terminal linkage places CR1 further away from TAZ2 than CR2.

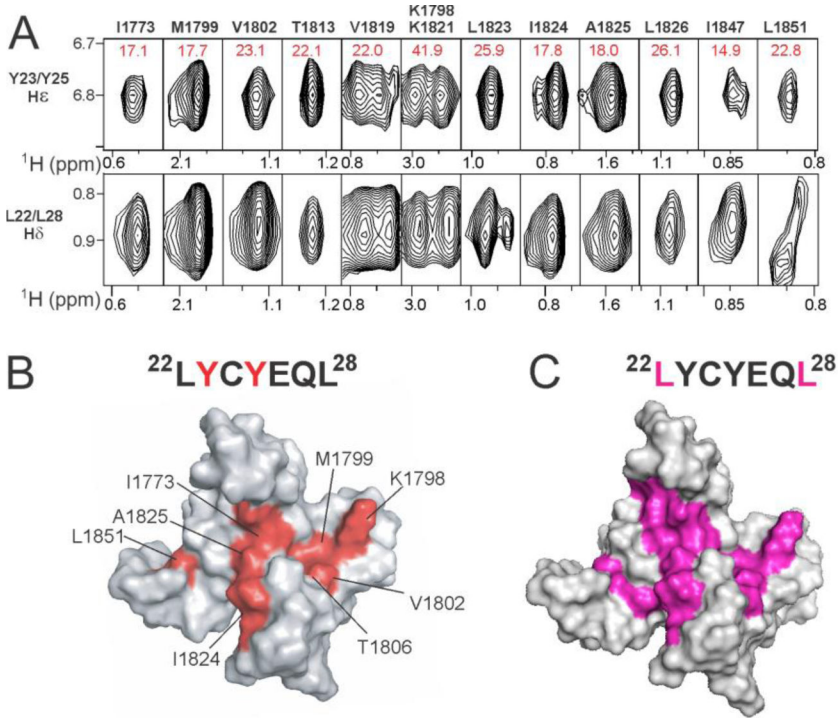


Figure 9.

Intermolecular contacts between TAZ2 and ppE7(17–51). (A) Intermolecular NOE cross peaks (recorded with a ^{13}C-filter-^{12}C-edited NOESY) between ^{13}C-TAZ2 in complex with ppE7(17–51) at a ratio of 1:1.2 and a mixing time of 200 ms. Red numbers in each panel refer to the ^{13}C chemical shift. A comprehensive version of this figure appears in Figure S6. (B) Illustration of the position on TAZ2 of residues with NOEs to one or both E7 CR2 tyrosines (red surface color). (C) Illustration of the position on TAZ2 of residues with NOEs to one or both E7 CR2 leucines (magenta surface color).

Table 1

Dissociation constants for E7 peptides from competition and NMR experiments

	K_D (μM)	K_D (μM)
	Fluorescence anisotropy	**NMR Spectroscopy**
E7(1–51)	0.10 ± 0.01	ND
E7(17–51)	1.8 ± 0.1	ND
CR2(18–37)	4.6 ± 0.2	3.4 ± 0.5
CR1(1–18)	31 ± 2	69 ± 7

The K_D values for each peptide are the average of four independent fluorescence anisotropy titrations. Errors were estimated by taking into account the uncertainty in the K_D of E7(1–51) determined by direct titration (K_D(direct) = 70 ± 14 nM).

Uncertainties in the NMR K_D are from uncertainties in fitting the titration data.