

Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F–DP

Ning Zheng,^{1,4} Ernest Fraenkel,^{2,3,4} Carl O. Pabo,² and Nikola P. Pavletich^{1,5}

¹Howard Hughes Medical Institute, Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021 USA; ²Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

The E2F and DP protein families form heterodimeric transcription factors that play a central role in the expression of cell cycle-regulated genes. The crystal structure of an E2F4–DP2–DNA complex shows that the DNA-binding domains of the E2F and DP proteins both have a fold related to the winged-helix DNA-binding motif. Recognition of the central c/gGCGCg/c sequence of the consensus DNA-binding site is symmetric, and amino acids that contact these bases are conserved among all known E2F and DP proteins. The asymmetry in the extended binding site TTTc/gGCGCc/g is associated with an amino-terminal extension of E2F4, in which an arginine binds in the minor groove near the TTT stretch. This arginine is invariant among E2Fs but not present in DPs. E2F4 and DP2 interact through an extensive protein–protein interface, and structural features of this interface suggest it contributes to the preference for heterodimers over homodimers in DNA binding.

[Key Words: E2F; DP; Winged-helix, DNA-binding domain; transcription factor; cell cycle]

Received January 12, 1999; revised version accepted January 22, 1999.

Progression through the G₁ and S phases of the eukaryotic cell cycle is tightly coupled to the transcriptional control of genes involved in growth and in DNA replication (Dynlacht 1997). In mammalian and many other eukaryotic cells, this temporal control of gene expression is carried out primarily by the E2F family of transcription factors (for review, see Slansky and Farnham 1996; Helin 1998). E2F-responsive genes include *c-myc* (Hiebert et al. 1989; Thalmeier et al. 1989), *cyclin A* (Schulze et al. 1995) and *cdc2* (Dalton 1992; Furukawa et al. 1994), growth-regulatory proteins, and dihydrofolate reductase (Blake and Azizkhan 1989; Mudryj et al. 1990), thymidine kinase (Dou et al. 1992) and DNA polymerase α (Pearson et al. 1991; DeGregori et al. 1995) proteins needed for DNA synthesis. E2F activity is controlled by the cell cycle machinery primarily through the binding of the retinoblastoma (Rb) family of pocket proteins (Chellappan et al. 1991; Chittenden et al. 1991; Cao et al. 1992; Shirodkar et al. 1992; Cobrinik et al. 1993; for review, see Dyson 1998), and through phosphorylation by the cyclin-dependent kinases (Dynlacht et al. 1994; Krek et al. 1994).

E2F proteins can, depending on whether they are associated with Rb, act either as repressors or as activators of transcription (Hiebert et al. 1992; Weintraub et al. 1992, 1995; Bremner et al. 1995). The Rb–E2F complexes, which predominate in quiescent or early G₁ cells (Bagchi et al. 1991; Cao et al. 1992), act as repressors of transcription as Rb masks the E2F transactivation domain (Helin et al. 1993a) and can also recruit a histone deacetylase at certain promoters (Brehm et al. 1998; Luo et al. 1998; Magnaghi-Jaulin et al. 1998). Free E2F transcription factors function as activators of transcription, and these are released following the phosphorylation of Rb at the G₁–S transition. Deletion of E2F1 in mice leads to atrophy of some tissues, and to dysplasia and tumors in other tissues, consistent with E2F1's dual role as an activator and a repressor of transcription in different contexts (Field et al. 1996; Yamasaki et al. 1996).

The E2F proteins form heterodimers with members of the DP family, which are distantly related to the E2F family. E2F proteins can also form homodimers, but heterodimerization with DPs enhances their DNA-binding, transactivation, and Rb-binding activities (Bandara et al. 1993; Helin et al. 1993b; Krek et al. 1993), and it appears that DP proteins are a component of all E2F activity in the cell (Wu et al. 1996). In humans, six E2F and two DP proteins have been isolated to date. E2Fs one through five share 20%–55% identity and have a similar organization of functional domains. E2F6 is divergent as it

³Present address: Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138 USA.

⁴These authors contributed equally to this work.

⁵Corresponding author.

E-MAIL nikola@xray2.mskcc.org; FAX (212) 717-3135.

lacks the transactivation domain and has been proposed to function as a repressor of E2F-dependent transcription (Morkel et al. 1997; Cartwright et al. 1998; Trimarchi et al. 1998). The two DP proteins are highly homologous (70%), and each can form functional heterodimers with any E2F family member (Wu et al. 1995).

Different E2F proteins are differentially regulated by members of the Rb family of pocket proteins. E2F1, E2F2, and E2F3 are bound and regulated by Rb (Lees et al. 1993), E2F5 binds to p130 only (Hijmans et al. 1995), whereas E2F4 associates with all three pocket proteins (Ikeda et al. 1996; Moberg et al. 1996). E2F proteins can also differ in the regulation of their subcellular localization. E2F1, E2F2, and E2F3 contain nuclear localization signals and are found exclusively in the nucleus. E2F4 and E2F5 appear to require other nuclear factors, such as their DP partners or the pocket proteins, to promote their nuclear localization (Muller et al. 1997; Verona et al. 1997).

Beside differences in their regulation, it is not yet entirely clear whether different E2F proteins differ in their DNA sequence specificity and in their preference for promoters of different E2F-responsive genes. On one hand, all E2F-DP combinations can bind to and transactivate at the consensus TTTc/gGCGCc/g E2F site (Lees et al. 1993; Buck et al. 1995; Zhang and Chellappan 1995). On the other hand, in vitro binding-site selection experiments have suggested that E2F-1 and E2F-4 may have differences in their specificity for variants of the E2F consensus site (Tao et al. 1997).

The E2F proteins contain an ~70-residue domain responsible for DNA binding (Kaelin et al. 1992; Cress et al. 1993; Ivey-Hoyle et al. 1993; O'Connor and Hearing 1994). DP proteins need a larger, 90-residue region for binding to DNA. The E2F and DP DNA-binding domains are followed by a hydrophobic heptad repeat involved in homo- and heterodimerization (Helin et al. 1993b). The E2F family also contains regions involved in transactivation and Rb binding (Kaelin et al. 1992; Cress et al. 1993). E2F and DP proteins share sequence homology in the last 30 residues of their DNA-binding domains and within their hydrophobic heptad repeat (Girling et al. 1993).

The hydrophobic heptad repeat is the primary dimerization domain, being necessary for the formation of E2F homodimers and E2F-DP heterodimers in the absence of DNA (Helin et al. 1993b). But several lines of evidence have suggested that the DNA-binding domains contribute to dimerization. The isolated DNA-binding domain of E2F1 binds DNA weakly (Krek et al. 1993; Jordan et al. 1994) and the corresponding region of DP1 does not bind detectably, but mixing the two DNA-binding domains greatly enhances DNA binding (Bandara et al. 1993; Fraenkel 1998). This has suggested that the DNA-binding domains may interact on the DNA and that this may contribute to dimerization and preference for DNA binding as heterodimers.

Here we report the 2.6-Å crystal structure of a complex containing the DNA-binding domains of E2F4 and DP2 bound to a DNA site from the adenovirus E2 promoter.

The structure reveals that the DNA-binding domains of E2F4 and DP2 adopt the winged-helix fold (Clark et al. 1993) and use residues invariant within their respective families to contact the bases of the DNA. This indicates that the winged-helix domains of other E2F-DP combinations will have very similar DNA sequence specificity. The structure also shows that E2F4 and DP2 form an extensive protein-protein interface, and structural features of this interface suggest that the winged-helix domains contribute to the preference of E2F and DP proteins to bind to DNA as heterodimers.

Results

Structures of the DNA-binding domains of E2F4 and DP2

The boundaries of the DNA-binding domains of E2F4 (residues 11–86) and DP2 (residues 60–154) used in the crystallization were identified by the proteolytic digestion of an E2F-DP heterodimer-DNA complex (Fig. 1; Fraenkel 1998). The heterodimer of the E2F4 and DP2

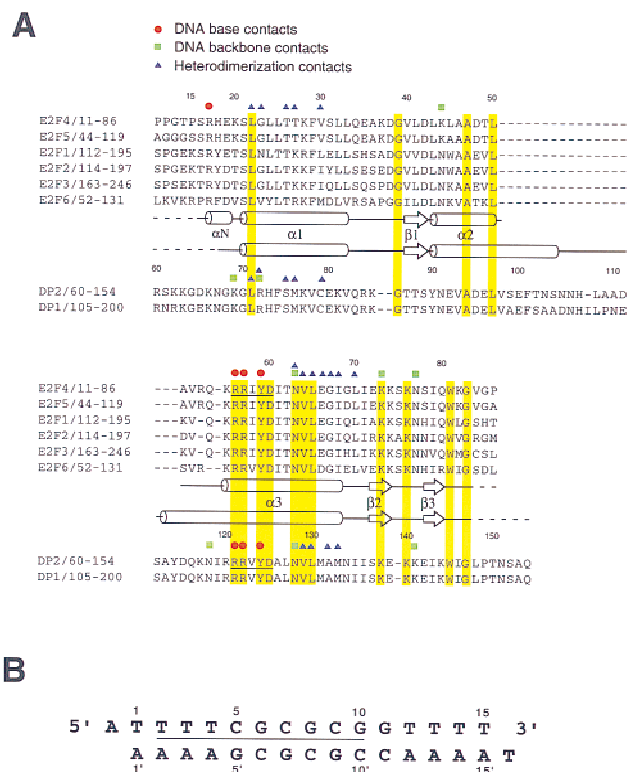


Figure 1. The DNA-binding domains of E2Fs and DPs have limited cross-family homology but share the same fold. (A) Sequence alignment of known E2F and DP family members with the E2F4 and DP2 polypeptides used in the crystallization. Residues conserved throughout the E2F and DP families are highlighted in yellow. The RRXYD DNA recognition motif is underlined in the E2F4 and DP2 sequences. (B) Sequence of the DNA duplex used in the cocrystallization, with the E2F site at the adenovirus E2 promoter underlined.

DNA-binding domains gives half-maximal DNA binding when each protein is at a concentration of 50 nM, compared with ~500 nM and >2000 nM for homodimers of E2F4 and DP2, respectively (data not shown).

The crystal structure, shown in Figure 2, reveals that the E2F DNA-binding domain has a structure related to the winged-helix DNA-binding motif, and not the basic helix-loop-helix (bHLH) motif that has been suggested (Kaelin et al. 1992; Cress et al. 1993). The winged-helix motif occurs in several eukaryotic transcription factors (for review, see Kaufmann and Knochel 1996), including HNF-3 γ (Clark et al. 1993), ETS1 (Kodandapani et al. 1996), and HSF (Harrison et al. 1994), as well as in the histone H5 globular domain (Ramakrishnan et al. 1993) (Fig. 3). This motif consists of three α helices and a β sheet, each contributing to a compact hydrophobic core. DP2 has the same overall structure as E2F4, except that the $\alpha 2$ and $\alpha 3$ helices of DP2 are longer by about two turns each, and E2F4 has an amino-terminal helical extension (αN ; Figs. 1A and 3A) that is not present in DP2. The remaining regions of E2F4 and DP2 superimpose quite well, with a 1.4-Å root mean square deviation (rmsd) for 59 residues (Fig. 3A). The structure-based alignment of E2F4 and DP2 shows that the regions corresponding to the $\alpha 1$ and $\alpha 2$ helices have only 8% sequence identity (Fig. 1A), and this, coupled to the greater length of the $\alpha 2$ and $\alpha 3$ helices of DP2, presumably complicated previous efforts to align the two families. The 30-residue region of clear homology between the E2F and DP families, termed the DEF box (Girling et al. 1993), coincides with the $\alpha 3$ helix and the $\beta 2$ and $\beta 3$ strands (Fig. 1A).

Compared with the winged-helix domain of HNF-3 γ (Clark et al. 1993), the E2F4 and DP2 domains do not contain the carboxy-terminal wing extension. DP2 and HNF-3 γ can be superimposed with an rmsd of 1.7 Å in the C α positions of 36 residues. E2F4 is more divergent, superimposing on HNF-3 γ with a 2.1-Å rmsd for 36 residues.

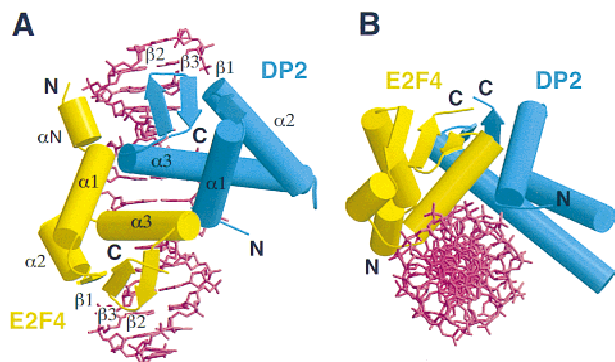


Figure 2. Structure of the E2F4-DP2 heterodimer DNA complex. (A) Schematic view looking down the approximate axis of twofold pseudo symmetry in the heterodimer. The DNA axis is vertical in this view. (B) View of the complex looking down the DNA axis. Figures were prepared with the programs MOLSCRIPT (Kraulis 1991) and RASTER3D (Merritt and Bacon 1997).

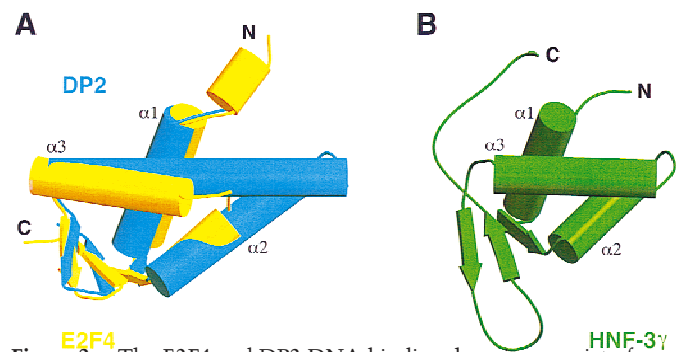


Figure 3. The E2F4 and DP2 DNA-binding domains consist of the winged-helix motif. (A) Superposition of the E2F4 and DP2 winged-helix DNA-binding domains. (B) The winged-helix domain of the HNF-3 γ transcription factor (Clark et al. 1993) in an orientation obtained by aligning it with the DP2 winged-helix domain in A.

DNA sequence recognition

The structure shows that the $\alpha 3$ helices of E2F4 and DP2 bind in the major groove of the DNA and make critical contacts to the edges of the bases. In both cases, the amino-termini of the $\alpha 1$ helices and portions of the β sheets contact the phosphodiester backbone of the DNA (Fig. 2). This overall DNA-binding arrangement for each protein is analogous to the other winged-helix proteins. However, winged-helix proteins typically bind DNA as monomers (Clark et al. 1993; Harrison et al. 1994; Kodandapani et al. 1996), whereas E2F4 and DP2 form an extensive hetero-oligomerization interface and present a continuous protein surface to the DNA (Fig. 2).

The E2F4-DP2 heterodimer binds to and recognizes the palindromic CGCGCG sequence of the binding site in an essentially symmetric arrangement (Fig. 4A,B). In this region, all of the DNA base contacts and most of the phosphate contacts are made by residues conserved throughout the E2F and DP families (Fig. 1A). E2F4 and DP2 each contact half of the palindromic sequence (CGCGCG and CGCGCG, top strand in Fig. 1B) by use of a conserved Arg-Arg-Xxx-Tyr-Asp sequence on their $\alpha 3$ helices (RRXYD motif, underlined in Fig. 1A). The binding of the E2F4 and DP2 $\alpha 3$ helices in the major groove is highly analogous, but a detailed comparison reveals that their positioning relative to the bases they contact differs by about 1 Å along the DNA axis.

The two arginines of the RRXYD motif each make a pair of hydrogen bonds with a guanine. These guanines occur in neighboring basepairs and are on opposite strands of the half site (CGC). The aspartic acid of the RRXYD motif makes charged hydrogen bonds to both arginines and appears to stabilize this arrangement (Fig. 4C). This pattern of hydrogen bonds is identical in the two halves of the heterodimer, except that DP2 has an additional residue, Asn118, helping to stabilize the arginines. The DP2 Asn118 side chain bridges the phosphodiester backbone of the DNA with the Arg122 guanidinium group (RRXYD, Fig. 4A,B). Several of the reported E2F-binding sites contain a half-site variant that has the c/gC sequence of the consensus half site re-

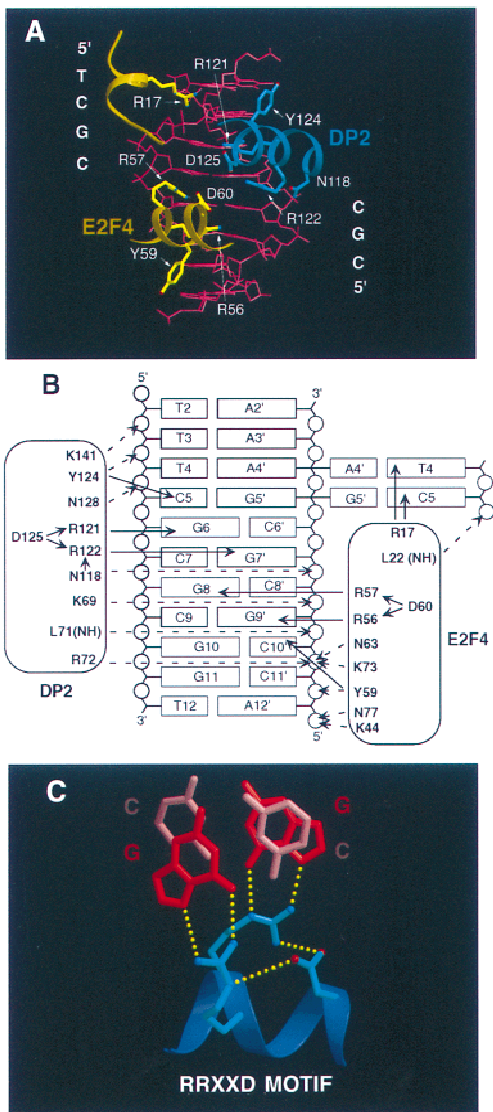


Figure 4. Recognition of the core DNA site by the E2F4–DP2 heterodimer is overall symmetric. (A) Figure shows the E2F4 and DP2 residues that contact the DNA bases; for clarity, only part of the $\alpha 3$ helices of each subunit and the amino-terminal helix of E2F4 are shown. (B) Sketch summarizes the DNA contacts made by the E2F4–DP2 heterodimer. (C) Close-up view of the interactions between the two arginines of the RRXXD motif and the two neighboring guanines within the half site.

placed with $c/g\overline{CC}$ (Slansky and Farnham 1996). Accordingly, we find that the mixture of the E2F4 and DP2 winged-helix domains has an affinity for a TTTCCC-GCG site that is comparable with the affinity for the TTTCGCGCG site used in the crystallization (data not shown). The crystal structure suggests that the arginine (RRXYD) that contacts this position could, in principle, reach to a guanine on either strand (Fig. 4C).

The outer position of the 3-bp half site (\overline{CGC}) is recognized by the tyrosine of the conserved RRXYD motif (Fig. 4A,B). The tyrosine phenyl group makes multiple van der Waals contacts to the C5 and C6 atoms of the

cytosine, whereas the side-chain hydroxyl group contacts phosphate and sugar groups. The close contacts between the tyrosine and the cytosine, which are very similar in the E2F and DP half sites, are consistent with sequence discrimination at this position suggested by the $c/g\overline{GC}$ consensus.

The consensus-binding site is not completely symmetric, having a T-rich portion at one end (TTTc/gGCGCg/c). In the crystal structure, this is associated with an E2F4 amino-terminal extension (residues 16–19) that is conserved in the E2F family but not in the DP family (Fig. 1A). This region forms the short αN helix and inserts an arginine side chain (Arg17) deep inside the minor groove of the DNA near the T-rich portion (Fig. 4A,B). The Arg-17 side chain contacts the O2 group of Thy4 and the O2 and sugar groups of Cyt5, and these contacts are associated with a compression of the minor groove in this region. On the basis of the electron density maps, Arg-17 is not as rigid as some of the other side-chain basepair contacts in the complex. However, the arginine contacts appear to be important because this residue is invariant among E2F family members (Fig. 1A) and its deletion abolishes DNA binding (Jordan et al. 1994). The corresponding region of DP2 (residues 60–68), which lacks the arginine, is disordered in our crystal structure and does not appear to be making direct DNA contacts. Similar arginine-minor groove contacts, associated with an A/T rich sequence and with a compressed minor groove have been observed in other protein–DNA complexes (Aggarwal et al. 1988; Clark et al. 1993; Cho et al. 1994).

The heterodimerization interface

The structure of the complex reveals an extensive protein–protein interface between the E2F4 and DP2 winged-helix domains (Figs. 2 and 5). This heterodimer interface is predominantly hydrophobic and buries a total of 1160 \AA^2 surface area. The interface contains the $\alpha 1$ and $\alpha 3$ helices from both E2F4 and DP2. These helices pack in triplets. The $\alpha 3$ helix of E2F4 packs in between the $\alpha 1$ and $\alpha 3$ helices of DP2 at $\sim 45^\circ$ and $\sim 90^\circ$, respectively, and the $\alpha 3$ helix of DP2 packs in a reciprocal fashion with the $\alpha 1$ and $\alpha 3$ helices of DP2. The arrangement of the helices at the interface is thus approximately symmetrical, explaining how the DNA-binding domain of E2F can also form homodimers (Huber et al. 1993). The E2F4 and DP2 residues at the interface are well conserved within their separate families, being 75% and 100% identical in the respective families. This suggests that other E2F and DP combinations will have similar heterodimer interfaces (Fig. 1A).

The symmetry at the heterodimer interface is less precise than that at the protein–DNA interface, owing to modest differences in the intermolecular packing of the E2F4 and DP2 structural elements (Fig. 5). In particular, contacts between the E2F4 $\alpha 3$ helix and the DP2 $\alpha 1$ helix are more extensive than the reciprocal contacts between the DP2 $\alpha 3$ and E2F4 $\alpha 1$ helices. The E2F4 $\alpha 3$ and DP2 $\alpha 1$ helices make 70 intermolecular van der Waals con-

Zheng et al.

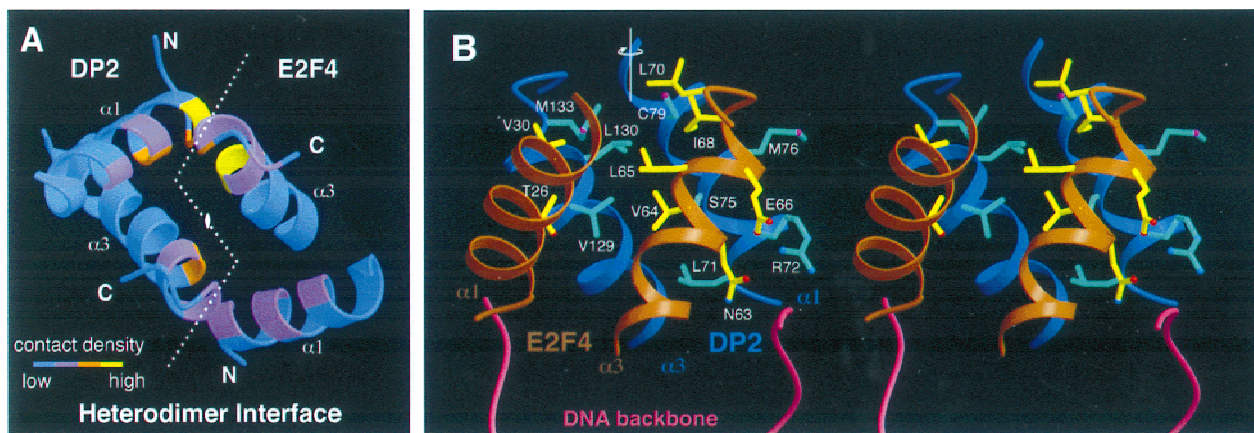


Figure 5. The interface of the E2F4–DP2 heterodimer involves the $\alpha 1$ and $\alpha 3$ helices of both proteins. (A) The interface has an approximate twofold symmetry, with the axis of symmetry (indicated) perpendicular to the plane of the figure. There are many differences in the intermolecular contacts, with the DP2 $\alpha 1$ and E2F4 $\alpha 3$ helices packing more extensively than the reciprocal E2F4 $\alpha 1$ and DP2 $\alpha 3$ helices. The intermolecular contact density was calculated by considering contacts made by each amino acid with an interatomic distance $< 4 \text{ \AA}$. (B) Stereo view of the E2F4–DP2 interface. To make the interface easier to see, only residues that make multiple van der Waals contacts ($< 4.0 \text{ \AA}$) are illustrated. A salt bridge is formed between the $\alpha 1$ helix of DP2 and $\alpha 3$ helix of E2F, contributing to the asymmetry at the interface.

tacts (considering interatomic distances $< 4.0 \text{ \AA}$), whereas the DP2 $\alpha 3$ and E2F4 $\alpha 1$ make only 20 such contacts (Fig. 5A). For example, the DP2 Leu-71 and E2F4 Val-64 side chains at the tightly packed half of the interface make a total of 21 intermolecular contacts with each other and with other interface residues; the equivalent residues at the loosely packed portion of the interface, the E2F4 Leu-22 and DP2 Val-129, make only one and six intermolecular contacts, respectively.

This asymmetry at the dimerization interface is associated with differences in the overall structures of DP2 and E2F4. As can be seen in Figure 3A, portions of the $\alpha 1$ and $\alpha 3$ helices of DP2 are closer together than the corresponding ones of E2F4 by $\sim 3 \text{ \AA}$. The closer $\alpha 1$ – $\alpha 3$ interhelical distance in DP2 is associated with small hydrophobic core residues, such as Val-78 and Ala-126, packing against each other at the interface of these two helices. In contrast, E2F4 generally has larger amino acids involved in the $\alpha 1$ – $\alpha 3$ helix packing. For example, the E2F4 residues corresponding to the Val-78 and Ala-126 pair of DP2 are Phe-29 and Ile-61, respectively. This difference in the DP2 and E2F4 structures is likely to be preserved in the structures of other family members because most of the residues in this portion of the hydrophobic core are well conserved within the separate families. The aforementioned Val-78 and Ala-126 residues of DP2 and Phe-29 and Ile-61 residues of E2F4 are invariant within their respective families. As a result of this difference between the structures of E2F4 and DP2, the protein–protein interfaces in the homodimers would necessarily be different from the interface observed in the heterodimer.

In addition to family-specific structural features, differences in the E2F4 and DP2 residues at the interface further contribute to the asymmetry in contacts (about two thirds of these residues differ between E2F4 and DP2). For example, Arg-72 of DP2 and Glu-66 of E2F4,

both conserved within their families, form a salt bridge in the densely packed half of the interface (Fig. 5B). In the sparsely packed half of the interface, the corresponding residues are Gly-72 of E2F4 and Met-131 of DP2, which make no contacts at all. The Arg–Glu salt bridge would not be possible in any homodimer (Fig. 1A).

Discussion

Our structure shows that the E2F and DP DNA-binding domains have the same overall fold, including the regions in which they share no significant homology. This fold is related to the winged-helix DNA-binding motif, except that E2F4 and DP2 lack a carboxy-terminal wing region present in other winged-helix proteins like HNF-3 γ (Clark et al. 1993). There are several structural differences between E2F4 and DP2. E2F4 contains an amino-terminal helical extension to the winged-helix motif that is absent from DP2 and HNF-3 γ , and also has a slightly different arrangement of its helices compared with those of DP2 and HNF-3 γ . In these respects, the E2F4 winged-helix domain is more divergent than that of DP2. On the other hand, DP2 contains an insertion between its $\alpha 2$ and $\alpha 3$ helices that results in the extension of these two helices by approximately two turns each compared with those of E2F4 and HNF-3 γ .

Despite these differences, E2F4 and DP2 make very similar contacts to the bases in the major groove of the DNA, recognizing the central CGCGCG sequence in an essentially symmetric fashion. The asymmetry in the extended TTTGCGCG-binding site is associated with the amino-terminal extension of E2F4 binding near the TTT sequence. This asymmetry in the contacts could help orient the E2F–DP heterodimer on the promoter. All of the E2F4 and DP2 residues that contact the bases and most of the residues that contact the phosphodiester backbone of the DNA are invariant in their respective

families. This suggests that other combinations of E2F and DP winged-helix domains would make very similar contacts to the 9-bp binding site. It is conceivable, however, that the sequence specificity of intact E2F–DP heterodimers may be modulated by residues outside of the winged-helix domains, or by other proteins bound to the E2F–DP complex making contacts to bases peripheral to the core binding site.

The E2F4 and DP2 winged-helix domains do not include the hydrophobic heptad repeat, which is necessary for homo and heterodimerization in the absence of DNA (Helin et al. 1993b). Nevertheless, they show the preference seen with the intact proteins to bind DNA as heterodimers instead of homodimers (Bandara et al. 1993; Helin et al. 1993b). The structure shows that the E2F4 and DP2 winged-helix domains form an extensive protein–protein interface. The arrangement of E2F4 and DP2 structural elements and their interactions at this interface have significant asymmetry, and this would contribute to the preference to bind DNA as a heterodimer.

Materials and methods

Protein purification

The human E2F4 (residues 11–86) and DP2 (residues 60–154) polypeptides were overexpressed as GST fusion proteins in *Escherichia coli* and isolated from the soluble cell lysate by glutathione affinity chromatography. Following cleavage from the GST by thrombin, they were purified by cation exchange and gel filtration chromatography and were concentrated by ultrafiltration. The purified E2F4 (0.6 mM) and DP2 (0.7 mM) polypeptides were first mixed together in 20 mM bis-Tris-propane-HCl (BTP), 150 mM NaCl, 5 mM dithiothreitol (DTT) at pH 7.5, and were then added to a solution of the DNA duplex and am-

monium acetate. The final solution contained 0.3–0.6 mM of the ternary complex in a buffer of 20 mM BTP, 100 mM NaCl, 5 mM DTT, and 300 mM ammonium acetate.

Crystallization and data collection

Crystals were grown at 4°C by the hanging drop vapor diffusion method by mixing the complex with an equal volume of reservoir solution containing 27%–32% polyethylene glycol (PEG) 400, 50 mM NaCl, 100 mM 2-(N-Morpholino) ethanesulfonic acid (MES), 20 mM DTT at pH 6.0. The crystals form in space group P3₁21, with a = b = 101.3, c = 73.5 Å. Heavy atom soaks were performed in crystallization buffer lacking DTT with 10.0 mM (Au-1 in Table 1) or 1.0 mM (Au-2) of K₂Au(CN)₄. Additional derivatives were obtained by growing crystals with DNA that had 5-iodouracil or 5-iodocytosine substitutions. Diffraction data were collected at –170°C with crystals flash frozen in crystallization buffer containing 35% PEG400. The derivative data were collected with an R-Axis IV imaging plate detector mounted on a Rigaku 200HB generator, and the native data set was collected at the F1 beamline of the Cornell High Energy Synchrotron Source. Data were processed with the programs DENZO and SCALEPACK.

Structure determination and refinement

The structure was determined by the multiple isomorphous replacement (MIR) method (Table 1). Initial MIR phases, calculated with the program MLPHARE (Collaborative Computational Project 4 1994), had a mean figure of merit of 0.73–3.5 Å and were improved by solvent flattening with the program DM (Collaborative Computational Project 4 1994). The MIR maps had continuous electron density for most of the E2F4 and DP2 polypeptides and the DNA. A model of the complex was built into the MIR electron density maps with the program O (Jones et al. 1991) and was refined by simulated annealing with the program CNS (Brunger et al. 1998). Restrained temperature factor refinement and solvent correction was done with CNS. The

Table 1. Statistics from the crystallographic analysis

Data set	Native	Au-1	Au-2	IdU(1)	IdU(2)	IdU(3)	IdC(11')
Resolution (Å)	2.6	4.7	3.0	3.1	3.5	3.1	3.1
Observations	41452	7078	31983	33211	18682	37245	30606
Unique reflections	12901	2092	8526	7837	5361	8443	7783
Data coverage (%)	96.5	88.9	96.4	98.9	96.5	97.7	97.8
R _{sym} (%)	5.4	8.6	5.8	8.9	8.4	7.4	7.5
MIR analysis (20.0–3.5 Å)							
phasing power	—	2.2	1.2	1.3	1.8	1.2	1.1
R _{cullis}	—	0.60	0.81	0.77	0.67	0.78	0.81

Refinement statistics

Resolution (Å)	Reflections (F > 1σ)	Total atoms	Water atoms	R-factor (%)	R-free (%)	rmsd		
						bonds (Å)	angles (°)	B-factor (Å ²)
15–2.6	12057	1887	87	22.9	28.3	0.012	1.75	2.9

(R_{sym}) $\sum_h \sum_i |I_h - I_{h,i}| / \sum_h \sum_i I_{h,i}$ for the intensity (*I*) of *i* observations of reflection *h*. Phasing power- $\langle F_{\lambda,i} \rangle / E$, where $\langle F_{\lambda,i} \rangle$ is the root mean square heavy atom structure factor and *E* is the residual lack of closure error.

R_{cullis} is the mean residual lack of closure error divided by the dispersive difference. R-factor- $\sum |F_{\text{obs}} - F_{\text{calc}}| / \sum |F_{\text{obs}}|$, where *F*_{obs} and *F*_{calc} are the observed and calculated structure factors, respectively. (R-free) R-factor calculated using 5% of the reflection data chosen randomly and omitted from the start of refinement. (rmsd) Root mean square deviations from ideal geometry and root mean square variation in the B-factor of bonded atoms.

Zheng et al.

final model consists of residues 15–83 of E2F4, residues 68–147 of DP2, the central 15 bp of the DNA duplex, and 87 water molecules. The single base overhangs of the DNA, residues 11–14 and residues 84–86 at the ends of E2F4, and residues 60–66 and residues 148–154 of DP2 have no electron density and are disordered in the crystals.

Electrophoretic mobility shift assay

DNA gel mobility shift assays were performed with 1 nM a radiolabeled 50-bp DNA fragment containing the binding site, in the presence of 5 mg/ml of nonspecific competitor DNA (sheared calf thymus DNA) and 75 mM NaCl. The apparent dissociation constants of the complexes were typically obtained from a seven-point titration of protein concentration.

Acknowledgments

We thank S. Geromanos and H. Erdjument-Bromage of the Sloan-Kettering Microchemistry Facility for amino-terminal sequence and mass spectrometry analyses; the staff of the Cornell High Energy Synchrotron Source MacChess for help with data collection; David King for mass spectrometry analyses; and C. Murray for administrative help. This work was supported by the Howard Hughes Medical Institute, the National Institutes of Health, the Dewitt Wallace Foundation, and the Samuel and May Rudin Foundation. E.F. was a Howard Hughes Medical Institute predoctoral fellow. We thank S.C. Harrison for providing support to E.F. during the final stages of this project. Coordinates have been deposited with the Brookhaven Protein Data Bank.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked 'advertisement' in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Aggarwal, A., D. Rodgers, M. Drottar, M. Ptashne, and S. Harrison. 1988. Recognition of a DNA operator by the repressor of phase 434: A view at high resolution. *Science* **242**: 899–907.
- Bagchi, S., R. Weinmann, and P. Raychaudhuri. 1991. The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. *Cell* **65**: 1063–1072.
- Bandara, L.R., V.M. Buck, M. Zamanian, L.H. Johnston, and T.N.B. La. 1993. Functional synergy between DP-1 and E2F-1 in the cell cycle-regulating transcription factor DRTF1/E2F. *EMBO J.* **12**: 4317–4324.
- Blake, M.C. and J.C. Azizkhan. 1989. Transcription factor E2F is required for efficient expression of the hamster dihydrofolate reductase gene in vitro and in vivo. *Mol. Cell. Biol.* **9**: 4994–5002.
- Brehm, A., E.A. Miska, D.J. McCance, J.L. Reid, A.J. Bannister, and T. Kouzarides. 1998. Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**: 597–601.
- Bremner, R., B.L. Cohen, M. Sopta, P.A. Hamel, C.J. Ingles, B.L. Gallie, and R.A. Phillips. 1995. Direct transcriptional repression by pRB and its reversal by specific cyclins. *Mol. Cell. Biol.* **15**: 3256–3265.
- Brunger, A.T., P.D. Adams, G.M. Clore, W.L. Delano, P. Gros, R.W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, and G.L. Warren. 1998. Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54**: 905–921.
- Buck, V., K.E. Allen, T. Sorensen, A. Bybee, E.M. Hijmans, P.M. Voorhoeve, R. Bernards, and T.N.B. La. 1995. Molecular and functional characterisation of E2F-5, a new member of the E2F family. *Oncogene* **11**: 31–38.
- Cao, L., B. Faha, M. Dembski, L.H. Tsai, E. Harlow, and N. Dyson. 1992. Independent binding of the retinoblastoma protein and p107 to the transcription factor E2F. *Nature* **355**: 176–179.
- Cartwright, P., H. Muller, C. Wagener, K. Holm, and K. Helin. 1998. E2F-6: A novel member of the E2F family is an inhibitor of E2F-dependent transcription. *Oncogene* **17**: 611–623.
- Chellappan, S.P., S. Hiebert, M. Mudryj, J.M. Horowitz, and J.R. Nevins. 1991. The E2F transcription factor is a cellular target for the RB protein. *Cell* **65**: 1053–1061.
- Chittenden, T., D.M. Livingston, and W.G. Kaelin Jr. 1991. The T/E1A-binding domain of the retinoblastoma product can interact selectively with a sequence-specific DNA-binding protein. *Cell* **65**: 1073–1082.
- Cho, Y., S. Gorina, P. Jeffrey, and N. Pavletich. 1994. Crystal structure of a p53 tumor suppressor-DNA complex: Understanding tumorigenic mutations. *Science* **265**: 346–355.
- Clark, K., E. Halay, E. Lai, and S. Burley. 1993. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**: 412–420.
- Cobrinik, D., P. Whyte, D.S. Peeper, T. Jacks, and R.A. Weinberg. 1993. Cell cycle-specific association of E2F with the p130 E1A-binding protein. *Genes & Dev.* **7**: 2392–2404.
- Collaborative Computational Project. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.* **D50**.
- Cress, W.D., D.G. Johnson, and J.R. Nevins. 1993. A genetic analysis of the E2F1 gene distinguishes regulation by Rb, p107, and adenovirus E4. *Mol. Cell. Biol.* **13**: 6314–6325.
- Dalton, S. 1992. Cell cycle regulation of the human cdc2 gene. *EMBO J.* **11**: 1797–1804.
- DeGregori, J., T. Kowalik, and J.R. Nevins. 1995. Cellular targets for activation by the E2F1 transcription factor include DNA synthesis- and G1/S-regulatory genes. *Mol. Cell. Biol.* **15**: 4215–4124.
- Dou, Q.P., P.J. Markell, and A.B. Pardee. 1992. Thymidine kinase transcription is regulated at G1/S phase by a complex that contains retinoblastoma-like protein and a cdc2 kinase. *Proc. Natl. Acad. Sci.* **89**: 3256–3260.
- Dynlacht, B. 1997. Regulation of transcription by proteins that control the cell cycle. *Nature* **389**: 149–152.
- Dynlacht, B.D., O. Flores, J.A. Lees, and E. Harlow. 1994. Differential regulation of E2F transactivation by cyclin/cdk2 complexes. *Genes & Dev.* **8**: 1772–1786.
- Dyson, N. 1998. The regulation of E2F by pRB-family proteins. *Genes & Dev.* **12**: 2245–2262.
- Field, S.J., F.Y. Tsai, F. Kuo, A.M. Zubiaga, W.G. Kaelin Jr., D.M. Livingston, S.H. Orkin, and M.E. Greenberg. 1996. E2F-1 functions in mice to promote apoptosis and suppress proliferation. *Cell* **85**: 549–561.
- Fraenkel, E. 1998. "Side chain mobility and hydration in homeodomain-DNA recognition." Ph.D. thesis, Massachusetts Institute of Technology, Boston, MA.
- Furukawa, Y., Y. Terui, K. Sakoe, M. Ohta, and M. Saito. 1994. The role of cellular transcription factor E2F in the regulation of cdc2 mRNA expression and cell cycle control of human hematopoietic cells. *J. Biol. Chem.* **269**: 26249–26258.
- Girling, R., J.F. Partridge, L.R. Bandara, N. Burden, N.F. Totty, J.J. Hsuan, and T.N.B. La. 1993. A new component of the

- transcription factor DRTF1/E2F. *Nature* **365**: 468.
- Harrison, C., A. Bohm, and H. Nelson. 1994. Crystal structure of the DNA binding domain of the heat shock transcription factor. *Science* **263**: 224–227.
- Helin, K. 1998. Regulation of cell proliferation by the E2F transcription factors. *Curr. Opin. Genet. Dev.* **8**: 28–35.
- Helin, K., E. Harlow, and A. Fattaey. 1993a. Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein. *Mol. Cell. Biol.* **13**: 6501–6508.
- Helin, K., C.L. Wu, A.R. Fattaey, J.A. Lees, B.D. Dynlacht, C. Ngwu, and E. Harlow. 1993b. Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation. *Genes & Dev.* **7**: 1850–1861.
- Hiebert, S.W., M. Lipp, and J.R. Nevins. 1989. E1A-dependent trans-activation of the human MYC promoter is mediated by the E2F factor. *Proc. Natl. Acad. Sci.* **86**: 3594–3598.
- Hiebert, S.W., S.P. Chellappan, J.M. Horowitz, and J.R. Nevins. 1992. The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. *Genes & Dev.* **6**: 177–185.
- Hijmans, E.M., P.M. Voorhoeve, R.L. Beijersbergen, t.V.L.J. van, and R. Bernards. 1995. E2F-5, a new E2F family member that interacts with p130 in vivo. *Mol. Cell. Biol.* **15**: 3082–3089.
- Huber, H.E., G. Edwards, P.J. Goodhart, D.R. Patrick, P.S. Huang, M. Ivey-Hoyle, S.F. Barnett, A. Oliff, and D.C. Heimbros. 1993. Transcription factor E2F binds DNA as a heterodimer. *Proc. Natl. Acad. Sci.* **90**: 3525–3529.
- Ikeda, M.A., L. Jakoi, and J.R. Nevins. 1996. A unique role for the Rb protein in controlling E2F accumulation during cell growth and differentiation. *Proc. Natl. Acad. Sci.* **93**: 3215–3220.
- Ivey-Hoyle, M., R. Conroy, H.E. Huber, P.J. Goodhart, A. Oliff, and D.C. Heimbros. 1993. Cloning and characterization of E2F-2, a novel protein with the biochemical properties of transcription factor E2F. *Mol. Cell. Biol.* **13**: 7802–7812.
- Jones, T., J. Zou, S. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.* **A47**: 110–119.
- Jordan, K., A. Haas, T. Logan, and D. Hall. 1994. Detailed analysis of the basic domain of the E2F1 transcription factor indicate that it is unique among bHLH proteins. *Oncogene* **9**: 1177–1185.
- Kaelin, W.G., Jr., W. Krek, W.R. Sellers, J.A. DeCaprio, F. Ajchenbaum, C.S. Fuchs, T. Chittenden, Y. Li, P.J. Farnham, M.A. Blonar et al. 1992. Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F-like properties. *Cell* **70**: 351–364.
- Kaufmann, E. and W. Knochel. 1996. Five years on the wings of fork head. *Mech. Dev.* **57**: 3–20.
- Kodandapani, R., F. Pio, C. Ni, G. Piccially, M. Klemsz, S. McKercher, R. Maki, and K. Ely. 1996. A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature* **380**: 456–460.
- Kraulis, P. 1991. Molscript: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**: 946–950.
- Krek, W., D.M. Livingston, and S. Shirodkar. 1993. Binding to DNA and the retinoblastoma gene product promoted by complex formation of different E2F family members. *Science* **262**: 1557–1560.
- Krek, W., M.E. Ewen, S. Shirodkar, Z. Arany, W.G. Kaelin Jr., and D.M. Livingston. 1994. Negative regulation of the growth-promoting transcription factor E2F-1 by a stably bound cyclin A-dependent protein kinase. *Cell* **78**: 161–172.
- Lees, J.A., M. Saito, M. Vidal, M. Valentine, T. Look, E. Harlow, N. Dyson, and K. Helin. 1993. The retinoblastoma protein binds to a family of E2F transcription factors. *Mol. Cell. Biol.* **13**: 7813–7125.
- Luo, R.X., A.A. Postigo, and D.C. Dean. 1998. Rb interacts with histone deacetylase to repress transcription. *Cell* **92**: 463–473.
- Magnaghi-Jaulin, L., R. Groisman, I. Naguibneva, P. Robin, S. Lorain, J. Le Villain, F. Troalen, D. Trouche, and A. Harel-Bellan. 1998. Retinoblastoma protein represses transcription by recruiting a histone deacetylase. *Nature* **391**: 601–605.
- Merritt, E.A. and D.J. Bacon. 1997. Raster3D—Photorealistic molecular graphics. *Methods Enzymol.* **277**: 505–524.
- Moberg, K., M.A. Starz, and J.A. Lees. 1996. E2F-4 switches from p130 to p107 and pRB in response to cell cycle reentry. *Mol. Cell. Biol.* **16**: 1436–1449.
- Morkel, M., J. Wenkel, A.J. Bannister, T. Kouzarides, and C. Hagemeier. 1997. An E2F-like repressor of transcription. *Nature* **390**: 567–568.
- Mudryj, M., S.W. Hiebert, and J.R. Nevins. 1990. A role for the adenovirus inducible E2F transcription factor in a proliferation dependent signal transduction pathway. *EMBO J.* **9**: 2179–2184.
- Muller, H., M.C. Moroni, E. Vigo, B.O. Petersen, J. Bartek, and K. Helin. 1997. Induction of S-phase entry by E2F transcription factors depends on their nuclear localization. *Mol. Cell. Biol.* **17**: 5508–5520.
- O'Connor, R.J. and P. Hearing. 1994. Mutually exclusive interaction of the adenovirus E4-6/7 protein and the retinoblastoma gene product with internal domains of E2F-1 and DP-1. *J. Virol.* **68**: 6848–6862.
- Pearson, B.E., H.P. Nasheuer, and T.S. Wang. 1991. Human DNA polymerase alpha gene: Sequences controlling expression in cycling and serum-stimulated cells. *Mol. Cell. Biol.* **11**: 2081–2095.
- Ramakrishnan, V., J. Finch, V. Graziano, P. Lee, and R. Sweet. 1993. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362**: 219–223.
- Schulze, A., K. Zerfass, D. Spitkovsky, S. Middendorp, J. Berges, K. Helin, P. Jansen-Durr, and B. Henglein. 1995. Cell cycle regulation of the cyclin A gene promoter is mediated by a variant E2F site. *Proc. Natl. Acad. Sci.* **92**: 11264–11268.
- Shirodkar, S., M. Ewen, J.A. DeCaprio, J. Morgan, D.M. Livingston, and T. Chittenden. 1992. The transcription factor E2F interacts with the retinoblastoma product and a p107-cyclin A complex in a cell cycle-regulated manner. *Cell* **68**: 157–166.
- Slansky, J.E. and P.J. Farnham. 1996. Introduction to the E2F family: Protein structure and gene regulation. *Curr. Top. Microbiol. Immunol.* **208**: 1–30.
- Tao, Y., R.F. Kassatly, W.D. Cress, and J.M. Horowitz. 1997. Subunit composition determines E2F DNA-binding site specificity. *Mol. Cell. Biol.* **17**: 6994–7007.
- Thalmeier, K., H. Synovzik, R. Mertz, E.L. Winnacker, and M. Lipp. 1989. Nuclear factor E2F mediates basic transcription and trans-activation by E1a of the human MYC promoter. *Genes & Dev.* **3**: 527–536.
- Trimarchi, J.M., B. Fairchild, R. Verona, K. Moberg, N. Andon, and J.A. Lees. 1998. E2F-6, a member of the E2F family that can behave as a transcriptional repressor. *Proc. Natl. Acad. Sci.* **95**: 2850–2855.
- Verona, R., K. Moberg, S. Estes, M. Starz, J.P. Vernon, and J.A. Lees. 1997. E2F activity is regulated by cell cycle-dependent changes in subcellular localization. *Mol. Cell. Biol.* **17**: 7268–7282.
- Weintraub, S.J., C.A. Prater, and D.C. Dean. 1992. Retinoblas-

Zheng et al.

- toma protein switches the E2F site from positive to negative element. *Nature* **358**: 259–261.
- Weintraub, S.J., K.N. Chow, R.X. Luo, S.H. Zhang, S. He, and D.C. Dean. 1995. Mechanism of active transcriptional repression by the retinoblastoma protein. *Nature* **375**: 812–815.
- Wu, C.L., L.R. Zukerberg, C. Ngwu, E. Harlow, and J.A. Lees. 1995. In vivo association of E2F and DP family proteins. *Mol. Cell. Biol.* **15**: 2536–2546.
- Wu, C.L., M. Classon, N. Dyson, and E. Harlow. 1996. Expression of dominant-negative mutant DP-1 blocks cell cycle progression in G1. *Mol. Cell. Biol.* **16**: 3698–3706.
- Yamasaki, L., T. Jacks, R. Bronson, E. Goillot, E. Harlow, and N.J. Dyson. 1996. Tumor induction and tissue atrophy in mice lacking E2F-1. *Cell* **85**: 537–548.
- Zhang, Y. and S.P. Chellappan. 1995. Cloning and characterization of human DP2, a novel dimerization partner of E2F. *Oncogene* **10**: 2085–2093.



Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F–DP

Ning Zheng, Ernest Fraenkel, Carl O. Pabo, et al.

Genes Dev. 1999, **13**:

References

This article cites 67 articles, 33 of which can be accessed free at:
<http://genesdev.cshlp.org/content/13/6/666.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

An advertisement banner for Dharmacon Reagents and Horizon. On the left, it says 'Dharmacon Reagents' with the tagline 'Custom synthesis, RNAi, and CRISPR solutions'. In the center, the text 'Infinite Reliability' is displayed in large white font, with a 'More' button below it. On the right, the 'horizon' logo is shown with the tagline 'a PerkinElmer company'. The background features a colorful, abstract image of what appears to be a DNA double helix or a similar molecular structure.