DNA-BINDING SPECIFICITY OF THE E2F GENE FAMILY,
AND CLONING AND CHARACTERIZATION
OF A NOVEL FAMILY MEMBER

by

Brian A. Fairchild

B.S., Cellular and Molecular Biology
University of Michigan, 1992

Submitted to the Department of Biology
in partial fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biology

at the

Massachusetts Institute of Technology

February 1999

© 1999 Massachusetts Institute of Technology
All rights reserved

Signature of Author ................................................................. Brian A. Fairchild
Department of Biology
January 7, 1999

Certified by ................................................................. Jacqueline A. Lees
Assistant Professor of Biology
Thesis Supervisor

Accepted by ................................................................. Terry Orr-Weaver
Professor of Biology
Co-Chair, Department Committee on Graduate Students
For Susan, on whose love and support I rely daily.
ABSTRACT

The E2F transcription factors have been implicated in the regulation of cell cycle progression and apoptosis. The E2F proteins are differentially regulated, and have distinct functions in addition to any shared ones. This work describes the identification of a novel E2F protein, E2F-6. Unlike the previously characterized E2F proteins, E2F-6 lacks a transcriptional activation domain and does not associate with the tumor suppressor pRB or its homologs p107 or p130. E2F-6 must therefore function in ways different from those of the previously identified E2F proteins. The identification of FIP-1, a novel zinc-finger protein which specifically interacts with E2F-6, is also described.

The distinct functions of the E2F proteins are presumed to result from the differential regulation of E2F-responsive target genes by individual E2F proteins. This work explores the hypothesis that this specificity of target gene regulation could be mediated by differences in the preferred DNA binding sequences. A PCR-based selection to determine the preferred DNA binding sequences for different E2F complexes identified substantially similar sequences for each E2F complex tested. Furthermore, these sequences were able to bind to each complex with similar affinities. The sequences also showed very little differential specificity when used in transcriptional activation assays with different E2F proteins. Specificity in target gene activation by particular E2F proteins is therefore more likely to be mediated by differences in promoter context than by differences in the sequences of the E2F sites.
# TABLE OF CONTENTS

**Abstract** ........................................................................................................................................... 3

**Chapter I: Introduction**

A. Overview .............................................................................................................................................. 7

B. Adenovirus: a tool to identify mechanisms of cellular proliferation control .................. 8
   B1. E1A deregulates cellular proliferation ................................................................. 9
   B2. p300 and CBP are bound, and possibly inactivated, by E1A ..................... 12

C. E1A inactivates the tumor and apoptosis suppressor pRB .............................. 15
   C1. RB is frequently inactivated in human tumors ............................................. 15
   C2. RB participates in restriction point control .................................................. 16
   C3. Cyclin-dependent kinases hyperphosphorylate pRB at the restriction point... 17
   C4. RB suppresses apoptosis ............................................................................. 23
   C5. RB inactivation is insufficient to promote transformation .................... 24
   C6. E1A also binds the pRB homologs p107 and p130 ............................... 25
   C7. pRB suppresses growth through its "large pocket" .................................... 27

D. E2F function is antagonized by pRB, p107, and p130 .............................. 29
   D1. The E2F and DP gene families ................................................................. 30
   D2. pRB-mediated repression ......................................................................... 31
   D3. p107 and p130 also bind E2F and can repress transcription ................... 33

E. E2F promotes proliferation and apoptosis ................................................ 35
   E1. E2F promotes progression through G1 ................................................... 35
   E2. E2F promotes transformation and apoptosis ........................................... 37
   E3. E2F as an oncogene, and as a tumor suppressor ..................................... 39

F. E2F genes: specificity in function and in regulation ................................. 41
   F1. Individual E2F genes have specific functions ........................................... 41
   F2. The E2F genes are regulated in overlapping but distinct ways ............... 43

G. E2F genes: transcriptional targets ................................................................. 47
   G1. Some E2F targets are regulated by repression, some by activation ....... 49
   G2. Individual E2F genes may regulate specific subsets
       of transcriptional targets ............................................................................. 50

H. Summary .................................................................................................................. 51

I. References .................................................................................................................. 52
Chapter II: The individual E2F family members bind a common DNA recognition sequence

A. Abstract .............................................................................................................69
B. Introduction .........................................................................................................70
C. Results .................................................................................................................74
   C1. Production of recombinant proteins.............................................................74
   C2. E2F site selection ............................................................................................77
   C3. Relative affinities of E2F complexes for selected sites..................................85
   C4. Sequence variations do not confer E2F complex specificity in transactivation ....................................................................................................................85
D. Discussion .............................................................................................................90
E. Materials and Methods .....................................................................................94
F. Acknowledgements ............................................................................................97
G. References ...........................................................................................................98

Chapter III: E2F-6, a member of the E2F family that can behave as a transcriptional repressor

A. Abstract .............................................................................................................102
B. Introduction .........................................................................................................103
C. Results .................................................................................................................106
   C1. Isolation of cDNAs encoding an E2F family member ..................................106
   C2. E2F-6 is widely expressed in vivo ...............................................................107
   C3. E2F-6 displays low-affinity E2F DNA-binding activity ...............................113
   C4. E2F-6 is unable to activate the transcription of E2F-responsive genes ..........116
D. Discussion ..........................................................................................................120
E. Materials and Methods .....................................................................................123
F. Acknowledgements ............................................................................................126
G. References ...........................................................................................................127

Chapter IV: FIP-1, a zinc-finger protein, interacts specifically with E2F-6

A. Abstract .............................................................................................................130
B. Introduction .........................................................................................................131
C. Results .................................................................................................................135
   C1. Isolation of FIP-1 .........................................................................................135
   C2. Expression pattern of FIP-1 .........................................................................140
   C3. Identification of E2F-6 regions interacting with FIP-1 .................................143
Chapter V: Conclusion

A. Cloning and characterization of novel E2F genes ........................................ 161
   A1. What are the biological functions of E2F-6? ........................................ 161
   A2. Additional E2F genes must exist ........................................ 164

B. Toward an understanding of the mechanisms of E2F target specificity .......... 166

C. Broadening our understanding of E2F: tumor genetics and genetic screens .... 170
   C1. E2F and cancer genetics ............................................................... 170
   C2. Genetic screens could uncover additional E2F functions or regulators .... 173
   C3. Genetics, molecular biology and biochemistry in the study of E2F .......... 174

D. References ...................................................................................... 176

Acknowledgements ..................................................................................... 178
CHAPTER I

Introduction

A. Overview

The transcription factor E2F confers a G1/S-specific pattern of transcription on a variety of cellular genes. This is accomplished in part through the association of E2F with members of the pRB family of transcriptional repressors during early in the G1 phase of the cell cycle. This association may be critical to the well-documented tumor-suppressive and apoptosis-suppressive activities of pRB. Similarly, the adenoviral protein E1A dissociates E2F from pRB family members, and this function of the E1A protein contributes to its transforming and pro-apoptotic activities. This chapter discusses the biological and biochemical activities of E1A, pRB, and E2F, with special attention to the genetic and biochemical heterogeneity of the E2F transcription factor.
B. Adenovirus: a tool to identify mechanisms of cellular proliferation control

The regulation of cell division and of programmed cell death is critical to the growth, development and reproduction of multicellular organisms. Deregulated cell growth and division is the hallmark and perhaps the definition of cancer. Cancer cells also demonstrate distinct changes in their susceptibility to programmed cell death. The proper regulation of these processes is therefore a study of enormous medical interest. Some viruses induce cells to enter the DNA synthesis phase (S phase) of the cell cycle, which promotes replication of the viral DNA. In addition to promoting re-entry of the cell cycle and, hence, cellular proliferation, these viruses also manipulate the regulation of apoptotic control pathways to prevent the premature death of an infected cell. Viruses have therefore been used extensively by cancer biologists as tools to identify key components of the cellular regulatory pathways controlling proliferation and apoptosis. One such virus is the adenovirus.

The adenoviruses are a family of double-stranded DNA viruses that infect humans and other animals. They have been the subjects of intensive study, in part because of their role in human disease: of the more than forty human-specific adenoviruses described so far, some are associated with respiratory infections, others with ocular disease and still others with gastrointestinal infections (see Lucher, 1995 for a review). Of particular interest to cancer biologists, however, is the observation that adenoviral infection can induce cellular transformation. They are also known to evade mechanisms inducing programmed cell death (apoptosis) (reviewed in White, 1995). The mechanisms through which adenoviruses override the normal regulation of cellular growth, proliferation, and death are being carefully investigated, primarily as a tool to improve our understanding of those normal regulatory mechanisms and how tumor cells might escape them.

Adenoviruses can infect either dividing or quiescent cells. Upon infection by adenovirus, quiescent cells enter the DNA synthesis phase (S phase) of the cell cycle, with a concomitant induction of the expression of dihydrofolate reductase (DHFR), thymidine kinase, and other genes associated with cellular DNA replication. This induction of reentry of quiescent cells into the cell
cycle is believed to be mediated by the products of the adenoviral gene first expressed upon infection, the E1A gene.

**B1. E1A deregulates cellular proliferation**

The adenoviral E1A gene is expressed as at least five different alternatively spliced messenger RNAs (mRNAs), of which the two most abundant are the two largest, known as 12S and 13S (reviewed in Zantema and van der Eb, 1995). The 12S and 13S mRNAs of adenovirus type 5 (Ad5), one of the most studied species of adenovirus, encode proteins of 243 and 289 amino acids, respectively. The E1A proteins from the various species of adenovirus are homologous to each other. This homology is concentrated in three conserved domains, referred to as CR1, CR2, and CR3. The 12S mRNA encodes both CR1 and CR2 as well as intervening and flanking sequences; the 13S mRNA encodes a nearly identical gene product, differing only by the additional presence of CR3.

**E1A drives proliferation and apoptosis**

Adenoviral strains bearing small deletions in their genome have been generated to study the roles of particular genes in the life cycle of the virus. The E2A and E2B genes, for example, are required for the replication of the viral genome, whereas the E1B gene products are required to prevent apoptosis of the host cell during infection (reviewed in Philipson, 1995). Viruses that lack the E1A region are deficient in multiple biological activities. The E1A gene is the first viral gene to be expressed after infection; expression of the 13S mRNA is essential for efficient transcription of the remaining viral genes and, therefore, for the execution of the remainder of the viral life cycle. Internal deletions within the E1A gene have revealed that this function requires CR3, consistent with the ability of the 13S but not the 12S splice forms to mediate this transactivation (reviewed in Jones, 1995). As mentioned above, adenoviruses that lack a functional E1B region induce apoptosis upon infection; this apoptosis is induced by the E1A gene products (White, 1995, and references therein).
The E1A gene products are also involved in the stimulation of cell cycle progression of the host cell. Whereas normal viruses will rapidly induce the host cell to enter S phase, viruses that lack the E1A region are unable to do so (Howe et al., 1990). This does not seem to result from the requirement for this region in the transactivation of other viral genes. CR3, which is necessary and sufficient to induce the expression of the viral genes, is not required for S phase induction: the 12S mRNA species, which lacks CR3, is competent to promote the reentry of the cell into the cell cycle (Lillie et al., 1987; Zerler et al., 1987; Bellett et al., 1989; Howe et al., 1990), and expression of E1A by plasmid microinjection is sufficient to induce S phase in the absence of viral infection (Lillie et al., 1987).

In addition to promoting re-entry of the host cell into the cell cycle, E1A can also promote the transformation of primary cells. Adenoviral infection of cells that do not support the complete viral life cycle can lead to cellular transformation (reviewed in Lucher, 1995); transformation is rarely observed in other cell types, presumably as a result of cellular lysis at the end of the infection cycle. A combination of deletion studies and transfection experiments have determined that the E1A gene is capable of mediating cellular transformation (reviewed in Williams et al., 1995). Expression of E1A in the absence of viral infection causes focus formation in primary baby rat kidney (BRK) cells, although the foci degenerate as a result of E1A-induced apoptosis (White, 1995, and references therein). Coexpression of E1A with another gene that prevents E1A-induced apoptosis, such as E1B, bcl-2, or activated ras, is sufficient to induce complete and stable transformation of primary cells (Whyte et al., 1989; White, 1995).
The effects of E1A are mediated by specific domains of the protein

Two distinct regions of the E1A protein contribute to its ability to promote cell proliferation. The first is the amino-terminal region of the protein, including CR1 and a non-conserved region at the extreme amino-terminus, and the second is CR2. Mutation of either of these regions impairs the ability of E1A to induce DNA synthesis in most assays, and mutation of both regions completely abrogates S phase induction by E1A (Lillie et al., 1987; Zerler et al., 1987; Bellett et al., 1989; Howe et al., 1990). The same two regions are required for the transforming activity of E1A (Lillie et al., 1986; Moran et al., 1986; Lillie et al., 1987; Whyte et al., 1988b). The ability of E1A to induce S phase and its ability to induce cellular transformation are therefore genetically inseparable. Interestingly, the same regions are also required for E1A-induced apoptosis (Samuelson and Lowe, 1997).

The E1A proteins do not seem to possess an intrinsic DNA-binding activity. It has therefore long been assumed that the cellular effects of E1A expression in the absence of viral infection are mediated by protein-protein interactions. Consequently, E1A-associated proteins have been intensively studied in an attempt to determine the mechanisms through which E1A promotes cellular transformation and apoptosis. Immunoprecipitation of E1A from an adenovirus-transformed cell line revealed that a number of proteins can be found in a stable complex with E1A; the most abundant of these have molecular weights of approximately 105, 107 and 300 kD (Yee and Branton, 1985; Harlow et al., 1986). These proteins are cellular in origin: they are also present in a variety of cells that have not been exposed to adenoviral infection (Harlow et al., 1986). The 300 kD protein, known as p300, requires both CR1 and the amino-terminus of E1A for efficient binding to that protein (Whyte et al., 1989; Eckner et al., 1994). Efficient binding of E1A to p107 requires the presence of CR2; maximal binding to p105 also requires part of CR1 (Whyte et al., 1989). Given that these same regions of E1A are known to mediate its effects on cellular proliferation and apoptosis, much study has been focused on p300, p105 and p107 as potential effectors of E1A function.
B2. p300 and CBP are bound, and possibly inactivated, by E1A

The p300 gene was identified by screening an expression library with an antiserum to p300 protein purified through its association through E1A (Eckner et al., 1994). p300 binds to the amino-terminus of E1A; mutations in CR1 or in the extreme amino-terminus of E1A are capable of disrupting the interaction between E1A and p300, which would be consistent with a role for p300 in E1A-induced cell proliferation or apoptosis. The p300 protein is homologous to the CREB-binding protein CBP (Eckner et al., 1994; Lundblad et al., 1995). This homology includes the region used by p300 to bind to E1A, and in vitro studies have found that CBP and p300 bind to E1A with similar affinities (Lundblad et al., 1995). Thus, both p300 and CBP are bound by E1A and one or both may participate in the deregulation of cellular proliferation by E1A.

Like CBP, p300 also binds to many other transcription factors in the cell, suggesting a role as a transcriptional adaptor. Both have been reported to bind a wide variety of transcription factors and to have histone acetyltransferase (HAT) activity (Giles et al., 1998), indicating that these proteins probably also have a role in chromatin remodeling at promoters that are being actively transcribed. Studies on E1A-mediated repression of the SV40 enhancer suggest that p300 and/or CBP may have a required role in the activation of some promoters and that E1A may sequester or otherwise inactivate p300 and CBP, preventing this activation. In addition, they are consistent with p300 and/or CBP mediation of at least some of the in vivo effects of E1A.

p300, CBP, and cellular proliferation

The regions of E1A required for binding to p300 and CBP also mediate E1A's promotion of cellular proliferation and apoptosis. In fact, a point mutation abrogating the ability of E1A to bind to p300 in a co-immunoprecipitation assay, but not its ability to bind p105 or p107, also abrogates the transforming activity of E1A (Whyte et al., 1988b; Whyte et al., 1989). Furthermore, no E1A mutation has been described in which the transforming activity is retained but p300-binding is not, and all tested mutations in E1A that retain p300-binding activity also retain
the ability to promote S phase entry and apoptosis. Nevertheless, no direct evidence currently indicates a role for either p300 or CBP in either the promotion of S phase or of apoptosis.

Studies of the effects of E1A on p300-mediated transcriptional modulation suggest that E1A neutralizes p300 function, at least in some contexts, perhaps through sequestration of the protein (see above). If inactivation of p300 or CBP by E1A does indeed promote cellular transformation, their inactivation might also be expected at some frequency in naturally occurring human tumors. To date, there has also been one report of somatic mutation in p300 in a colorectal carcinoma and a gastric carcinoma; these are associated with loss of heterozygosity at the locus, suggesting that the mutations may be more likely to be hypomorphic than neomorphic (Muraoka et al., 1996). Patients with Rubinstein-Taybi syndrome, a disease caused by inheritance of an inactivated allele of the CBP gene, have been reported to bear a mildly increased predisposition to certain tumor types, particularly to certain tumors of the nervous system (Miller and Rubinstein, 1995). CBP may, therefore, act as a tumor suppressor in some cellular contexts, although no studies have yet addressed the fate of the second CBP allele in Rubinstein-Taybi tumors.

In addition, both p300 and CBP have been found in translocations with MLL, a homolog of the Drosophila Trithorax gene, in leukemia patients (reviewed in Giles et al., 1998). CBP has also been found in translocations with MOZ, a putative acetyltransferase, in patients with acute myelocytic leukemia (Borrow et al., 1996). The fusion products of these translocations are believed to be expressed, and are believed to be tumorigenic as a result of a gain-of-function, in contrast to the hypomorphic mutations observed in Rubenstein-Taybi and the presumably hypomorphic mutations in the two carcinomas mentioned above, suggesting that the mechanism for their tumorigenicity may be unrelated to the mechanism(s) in the loss-of-function scenarios.

In summary, the sequestration of p300 or CBP is required for some of the transcriptional effects of E1A. Sequestration or other modification of their activity might also play a role in the deregulation of cell proliferation or apoptosis by E1A, although no direct evidence supports this hypothesis. The alterations in p300 and CBP in some human tumors does suggest that these proteins may participate in the regulation of cellular proliferation in some contexts. As
transcriptional adaptors, however, both p300 and CBP have been reported to bind a wide variety of transcription factors. p300 and CBP are therefore potentially implicated in an immense number of signaling pathways. Perhaps only the discovery of a tumorigenic mutation specifically affecting only a limited subset of these pathways will permit cancer biologists to identify which functions of p300 and CBP are relevant to proliferation control and most urgently deserving of further study.
C. E1A inactivates the tumor and apoptosis suppressor pRB

Like p300, the 105 kD E1A-associated protein, p105, binds to regions of E1A that mediate its effects on cellular proliferation and apoptosis; in the case of p105, optimal binding to E1A requires both CR1 and CR2. A combination of immunological methods and protease digestion experiments revealed that p105 is identical to pRB, the product of the retinoblastoma-susceptibility gene, RB (Whyte et al., 1988a).

C1. RB is frequently inactivated in human tumors

The RB gene was originally identified based on its frequent mutation, often homozygous deletion, in human retinoblastomas (Friend et al., 1986; Friend et al., 1987; Lee et al., 1987). Predisposition to retinoblastomas is inherited as an autosomal dominant trait, and usually results in bilateral retinoblastomas, whereas unilateral retinoblastomas are more often the result of sporadic mutations (Knudson, 1971). Statistical analyses on patients with unilateral or bilateral retinoblastomas suggested a requirement for two mutational events with similar mutation rates in the genesis of retinoblastomas (Knudson, 1971). The frequent loss of heterozygosity observed in retinoblastomas on the long arm of chromosome 13 in the vicinity of the retinoblastoma susceptibility locus is consistent with a tumor-suppressive function for this locus (Cavenee et al., 1983; Dryja et al., 1984). Although gross genomic deletions of RB, the retinoblastoma-susceptibility gene, are found in only about 30% of retinoblastoma cell lines studied (Friend et al., 1986; Friend et al., 1987; Lee et al., 1987), gross alterations in the transcript are seen in almost all retinoblastoma cells (Friend et al., 1987; Lee et al., 1987). In one study the protein was undetectable in all eighteen retinoblastoma cultures tested, including thirteen short-term cultures (Horowitz et al., 1990). This locus is believed to be inactivated in all human retinoblastomas, and this inactivation is rate-limiting for tumor formation.

Humans who inherit a one defective (and one functional) copy of the RB gene develop retinoblastomas with approximately 95% penetrance, and most of these retinoblastomas are bilateral. Patients who survive retinoblastoma often develop other tumors, including sarcomas.
such as osteosarcomas. Indeed, mutations of the RB gene have been detected in primary tumors and/or tumor-derived cell lines from osteosarcomas, synovial sarcomas, small cell lung carcinomas, breast carcinomas, prostate carcinomas, cervical carcinomas and other tumors in addition to retinoblastomas (Harbour et al., 1988; Lee et al., 1988; T’Ang et al., 1988; Yokota et al., 1988; Bookstein et al., 1989; Varley et al., 1989; Horowitz et al., 1990). These mutations are often biallelic deletions, although missense and nonsense mutations have also been described. Thus, the RB gene clearly plays an important role in tumor suppression in a wide variety of human cell types.

A role for RB in the prevention of tumorigenesis is also supported by mouse studies. Mice that inherit an inactivated allele of RB develop pituitary tumors, and less frequently cancer of the thyroid (Jacks et al., 1992; Williams et al., 1994). The cause(s) of the differences in the tissue specificity of tumorigenesis between mice and humans in the case of RB mutation is (are) not known; nevertheless, the RB gene does play an essential role in the prevention of tumorigenesis in both species.

C2. RB participates in restriction point control

The commitment of a mammalian cell to replicate or to remain quiescent is usually determined during the G1 phase of the cell cycle. During early- to mid-G1 the growth of mammalian cells in tissue culture is dependent on the continuous presence of growth factors, usually provided by supplementation of the growth medium with serum; removal of the serum during this period of G1 induces the cells to enter a quiescent state (reviewed in Pardee, 1989). Later, during late G1, S or G2, progression through the cell cycle becomes serum-independent. The point at which the cell switches from serum-dependence to serum-independence is termed the restriction point (Pardee, 1989). The period of serum-dependence is also characterized by a marked sensitivity to the protein synthesis inhibitor cycloheximide: fibroblasts treated with cycloheximide will arrest in early G1, but once they have passed the restriction point they become much less sensitive to cycloheximide levels (reviewed in Pardee, 1989).
The RB gene contributes to control over passage through the restriction point. Mouse embryonic fibroblasts (MEFs) in which both alleles of the RB gene have been mutated lose some, but not all, aspects of restriction point control. RB-deficient MEFs are smaller than wild-type MEFs and spend less time in G1 (Herrera et al., 1996; Lukas et al., 1995). RB-deficient MEFs lack the extreme sensitivity of wild-type MEFs to cycloheximide during early- and mid-G1 (Herrera et al., 1996). Serum-dependence is also markedly reduced in RB-deficient MEFs (Lukas et al., 1995; Brugarolas et al., 1998), although not eliminated: complete serum withdrawal before late G1 will still arrest RB-deficient MEFs (Herrera et al., 1996). Given that RB-deficient MEFs retain some degree of serum-dependence, loss of RB does not abrogate all forms of restriction point control, but RB is nevertheless required for complete enforcement of the checkpoint.

RB mutation also leads to a relaxation of control over the G1/S transition in some tissues of the developing mouse. Specifically, aberrant S phase entry is observed in RB-deficient mice in the lens and both the central and peripheral nervous systems (Morgenbesser et al., 1994; Macleod et al., 1996). Inactivation of the RB gene is therefore sufficient to permit inappropriate entry into the cell cycle of some cell types that are normally quiescent. As noted above, the regions used by E1A to bind the RB-encoded protein, pRB, are involved in the promotion of S phase by E1A. Given that E1A expression and RB inactivation share this phenotype, binding of E1A to pRB may inactivate a pRB function and thus prevent (full) enforcement of normal controls on the G1/S transition.

C3. Cyclin-dependent kinases hyperphosphorylate pRB at the restriction point

Concomitant with passage of the cell through the restriction point, the majority of the RB protein, pRB, becomes hyperphosphorylated (DeCaprio et al., 1989). This hyperphosphorylation is mediated by the cyclin-dependent kinases (CDKs). CDKs have been demonstrated to play a central role in promoting progression through the cell cycle in a wide variety of organisms, including fusion and fission yeast, fruit flies, frogs, and mammals, among others. CDK activity
requires association of the CDK subunit with a cyclin subunit; modulation of cyclin protein levels during the cell cycle seems to be one mechanism regulating CDK activity during the cell cycle.

_Cyclins D and E promote passage through G1_

The first CDK complexes known to be active in the G1 phase of the mammalian cell cycle are CDK4 and CDK6 in association with one of the "D-type" cyclins, D1, D2, and D3. Treatment of mammalian cells in tissue culture with mitogens induces expression of D-type cyclins at the RNA and protein levels, but both disappear rapidly upon withdrawal of the mitogens (Matsushime et al., 1991). Overexpression of D-type cyclins leads to a shortening of G1 and smaller cell size (Jiang et al., 1993; Quelle et al., 1993; Resnitzky et al., 1994). Rodent fibroblast cells overexpressing D-type cyclins also show a reduced (but not eliminated) dependency on serum for cell cycle progression (Quelle et al., 1993; Resnitzky et al., 1994). Microinjection of a monoclonal antibody to cyclin D1 during early G1 prevents S phase entry, whereas no effect is observed closer to the G1/S transition (Baldin et al., 1993; Quelle et al., 1993). Similar results are observed for antibodies to cyclin D2 (Lukas et al., 1995b) or cyclin D3 (Bartkova et al., 1998) in cells expressing these cyclins. Thus, the D-type cyclins, presumably in combination with their partners CDK4 and CDK6, play an essential role in promoting passage through G1, and this role is arguably preceding or concomitant with passage through the restriction point.

Like the D-type cyclins in conjunction with CDK4 and CDK6, cyclin E/CDK2 plays an essential role in the G1/S transition. Ectopic expression of cyclin E shortens the G1 phase of the cell cycle; coexpression of cyclin D1 with cyclin E shortens G1 even further (Resnitzky et al., 1994; Resnitzky and Reed, 1995). Microinjection of an antibody to cyclin E during early G1 prevents passage through the G1/S transition, whereas microinjection later has no effect (Ohtsubo et al., 1995). Similarly, overexpression of dominant negative CDK2 arrests cells in G1 (Hofmann and Livingston, 1996).
Cyclins D and E participate in the hyperphosphorylation of pRB

Multiple lines of evidence indicate that pRB is phosphorylated by CDK complexes with D-type cyclin subunits. The D-type cyclins can bind directly to pRB (Dowdy et al., 1993; Ewen et al., 1993) and, in combination with CDK4 or CDK6, will phosphorylate pRB in vitro (Ewen et al., 1993; Kato and Sherr, 1993) on sites that are also phosphorylated in vivo (Kato and Sherr, 1993). Cyclin D/CDK4 and cyclin D/CDK6 complexes also phosphorylate pRB more efficiently than they phosphorylate histone H1; they are the only CDKs known to have this substrate specificity. Co-expression of cyclin D1 with CDK4 induces hyperphosphorylation of pRB in transient transfections (Lundberg and Weinberg, 1998). Furthermore, pRB is maintained in a primarily hypophosphorylated state in cells overexpressing p16, a specific inhibitor of CDK4 and CDK6, indicating that these kinase activities are required for full phosphorylation of the protein (Lundberg and Weinberg, 1998).

pRB is also phosphorylated by cyclin E/CDK2 complexes. Overexpression of cyclin E is sufficient to induce hyperphosphorylation of pRB (Hinds et al., 1992; Lundberg and Weinberg, 1998), and cyclin E/CDK2 complexes can phosphorylate pRB directly in vitro (Sherr, 1994). Overexpression of a dominant negative form of CDK2 prevents pRB from reaching a fully phosphorylated state (Lundberg and Weinberg, 1998). Cyclin E protein levels and cyclin E-associated kinase activity both peak in late G1 (Ohtsubo et al., 1995), at the time when pRB first becomes hyperphosphorylated. The phosphorylation of specific residues of pRB is significantly restricted in irradiated wild-type MEFs, in which the CDK inhibitor p21 is induced and cyclin E-associated kinase activity is inhibited, but not in p21-deficient MEFs, suggesting that specific pRB phosphorylation events require active cyclin E/CDK2 complexes (Moberg, 1998). The phosphorylation of pRB on serine 780, in contrast, is mediated specifically by cyclin D (Kitagawa et al., 1996).
Inactivation of pRB, in part by D-type cyclins, is required for passage through G1

Both D-type cyclins and cyclin E, together with their cognate CDK subunits, can therefore phosphorylate pRB \textit{in vitro}, are required for full pRB hyperphosphorylation \textit{in vivo}, and promote passage through the G1/S transition. In the case of the D-type cyclins, the inactivation of pRB seems to be the only function required for cell cycle progression. Microinjection of an antibody to cyclin D1 can only induce a G1 arrest in cells with wild-type pRB: cells without functional pRB are unaffected, although sensitivity can be restored by ectopic expression of wild-type pRB in these cells (Lukas et al., 1994; Lukas et al., 1995a). Similarly, only cells with wild-type pRB can be arrested in G1 by ectopic expression of the CDK4/CDK6 inhibitor p16 (Lukas et al., 1995).

Because overexpression of D-type cyclins and mutations inactivating pRB both promote S-phase progression and because inactivation of pRB relieves the requirement for D-type cyclin-associated kinase activity, it seems clear that 1) pRB is a critical target for this kinase activity \textit{in vivo} and 2) that this activity is required for the hyperphosphorylation and inactivation of pRB and for the subsequent passage through the G1/S transition. This must, furthermore, be the only requisite target for this activity in cells grown in tissue culture, given that the activity is dispensable in cells that lack wild-type pRB. Cyclin E, in contrast, must have additional critical targets \textit{in vivo}: microinjection of an antibody to cyclin E inhibits cell cycle progression even in cells that lack functional pRB (Ohtsubo et al., 1995).

Hyperphosphorylated pRB no longer interacts with a viral oncoprotein

It is interesting to note that E1A is not the only viral oncoprotein that is known to bind pRB. SV40 large T antigen and the E7 protein of human papilloma virus (HPV) are among the best known transforming viral proteins (other than E1A) that are known to interact with pRB through regions required for their transforming functions (DeCaprio et al., 1988; Dyson et al., 1989; Dyson et al., 1992). At least SV40 large T seems to interact specifically with the underphosphorylated form of pRB (DeCaprio et al., 1988), and the same is suspected to be true of the other viral oncoproteins (Weinberg, 1995). Hyperphosphorylated pRB has therefore lost at
least two functions: inhibition of S phase progression and binding to viral oncoproteins. This also indicates that any functions of pRB that must be abrogated by the oncoproteins are absent in the hyperphosphorylated protein. Oncoprotein binding and CDK-mediated hyperphosphorylation interfere with the same function of pRB, and this function is required for progression to S phase.

**Genes regulating the hyperphosphorylation of pRB are also mutated in human tumors**

This requirement for the inactivation of pRB prior to passage through the G1/S transition is likely to be the reason RB is often mutated in human tumors. If this is the case, extragenic mutations that also have the effect of inactivating pRB should also be observed in some tumors. Indeed, mutations that cause elevated levels of D-type cyclin-associated kinase activity are frequently observed in human tumors. Cyclin D1 was originally identified through its rearrangement in some B-cell lymphomas (Motokura et al., 1991), and cyclin D1 levels have also been shown to be elevated in a number of other tumor types, in some cases as a result of amplification of the locus (Herwig and Strauss, 1997). Cyclin D2 has also been shown to be a proto-oncogene in the mouse (Hanna et al., 1993). Amplification of CDK4 has also been observed in some gliomas and glioblastomas (He et al., 1994; Schmidt et al., 1994). Similarly, mutation of p16 is observed in familial melanoma (Hussussian et al., 1994; Kamb et al., 1994), and deletion of p16, often in combination with its homolog p15, another specific inhibitor of CDK4 and CDK6, is observed in a wide variety of tumor types (Weinberg, 1995). Although the interpretation of the deletion studies is complicated by the recent identification of another tumor suppressor, p19-ARF, present at the same locus (Quelle et al., 1995; Kamijo et al., 1997), the identification of mutations affecting only p16 prove that p16 is clearly a tumor suppressor in its own right (Quelle et al., 1997). Furthermore, the observation that small cell lung carcinomas almost invariably demonstrate mutations affecting either pRB expression or p16 expression, but not both, argues strongly that pRB and p16 are components of the same oft-targeted pathway in the tumorigenesis of these tissues (Otterson et al., 1994). Thus, the inactivation of pRB by CDK complexes promotes passage through the restriction point, and the mutagenesis of either pRB,
Figure 1. Inactivation of pRB is associated with tumorigenesis.

Inhibitory interactions are depicted by inverted "T" symbols.
CDK complexes, or CDK inhibitors can lead to a loss of proliferative control and to tumorigenesis (Figure 1).

C4. RB suppresses apoptosis

Some of the tissues in which unscheduled S phase entry is observed in RB-deficient mice are also prone to apoptosis (Morgenbesser et al., 1994; Macleod et al., 1996). The abnormal levels of apoptosis observed in the RB-deficient lens or central nervous system (CNS) are dependent on the presence of wild-type p53, a tumor suppressor and an inducer of apoptosis: mice that lack both pRB and p53 do not exhibit elevated levels of apoptosis in either the lens or CNS, although aberrant S phase entry is still observed (Morgenbesser et al., 1994; Macleod et al., 1996). Aberrant apoptotic figures are also observed in the peripheral nervous systems (PNS) of RB-deficient mice, but this apoptosis is apparently p53-independent (Macleod et al., 1996). RB-deficient MEFs are also sensitized to apoptosis-inducing drugs such as methotrexate (Almasan et al., 1995).

Overexpression studies also suggest pRB can act to inhibit apoptosis. HeLa cells are a human cervical carcinoma-derived cell line expressing low levels of HPV oncoproteins, inactivating their endogenous pRB and p53; overexpression of pRB can protect HeLa cells from apoptosis induced by overexpressed p53 (Haupt et al., 1995). Expression of wild-type pRB in SAOS-2 cells, a tumor cell line lacking functional pRB or p53, promotes cell survival following treatment with ionizing radiation (Haas-Kogan et al., 1995). Wild-type pRB can therefore inhibit either p53-dependent or -independent apoptosis in certain contexts.

The anti-apoptotic effects of pRB are consistent with the observed pro-apoptotic effects of E1A proteins that can bind and inactivate pRB. They are also consistent with the known pro-apoptotic effects of cyclin D1. Overexpression of cyclin D1 can induce apoptosis in multiple cell lines (Kranenburg et al., 1996; Sofer-Levi and Resnitzky, 1996). Elevated levels of cyclin D1 have been observed in neuronal cells undergoing apoptosis in response to growth factor withdrawal, suggesting that cyclin D1-mediated apoptosis may indeed have in vivo relevance.
Overexpression of the CDK4/6 inhibitor p16 has been shown to inhibit apoptosis during myocyte differentiation (Wang and Walsh, 1996). Thus, either the loss of pRB function through mutagenesis or its inactivation through hyperphosphorylation can promote apoptosis in a variety of tissues.

C5. RB inactivation is insufficient to promote transformation

pRB is a plausible target for E1A-mediated S phase entry and apoptosis and mutations that inactivate pRB promote tumorigenesis. In addition, inactivation of pRB may eventually prove to be required for cellular transformation in vitro. For example, mutants of SV40 large T antigen that have lost the ability to bind to pRB have also lost the ability to transform wild-type MEFs (Christensen and Imperiale, 1995; Zalvide and DeCaprio, 1995). There is, however, no evidence that the loss of pRB is sufficient to promote cellular transformation. pRB-deficient MEFs are not transformed; like wild-type MEFs, pRB-deficient MEFs require functional oncogenes to induce transformation. The large T mutants that cannot bind pRB are also unable to transform RB-deficient MEFs (Christensen and Imperiale, 1995; Zalvide and DeCaprio, 1995). Similar results were obtained with studies using the HPV E7 protein (Zalvide and DeCaprio, 1995). Thus, the inactivation of pRB does not, on its own, significantly increase the efficiency of cellular transformation. Furthermore, the pRB binding domain of the oncoprotein is required even in cells lacking pRB; inactivation of this region must also compromise additional functions required for transformation. In fact, the regions of E1A required for binding to pRB are also involved in binding to the E1A-associated proteins p107 and p130 (Whyte et al., 1989; Dyson et al., 1992), raising the possibility that they, alone or in combination with pRB, may mediate the effects of E1A on cellular transformation.

Indeed, although RB loss promotes tumorigenesis, S phase entry, and apoptosis, mouse embryos without a functional RB gene survive until embryonic day 13.5; death is most likely the result of a hematopoietic defect (Jacks et al., 1992). Although some tissues exhibit abnormally high levels of S phase entry and apoptosis, most embryonic tissues appear entirely normal, having
undergone many normal cell cycles and appropriate differentiation. In chimeric mouse embryos carrying some percentage of RB-deficient cells, those cells can contribute to most, if not all, tissues (Williams et al., 1994). In conclusion, despite the clear importance of pRB in preventing tumorigenesis and its contribution to the prevention of inappropriate S phase entry and apoptosis, there are clearly additional safeguards in most cells working in concert with the pRB pathway to prevent the deregulation of these processes.

C6. E1A also binds the pRB homologs p107 and p130

The E1A-associated proteins p107 and p130 have been cloned and are homologs of pRB (Ewen et al., 1991; Hannon et al., 1993; Li et al., 1993; Mayol et al., 1993). p107 and p130 are more homologous to each other than either is to pRB; nevertheless, they share noticeable homology to pRB throughout their length, particularly in the carboxy-terminal half of the protein. They also share an insertion between two domains highly conserved among pRB, p107, and p130. These two domains are referred to collectively as the "small pocket"; the insertion is referred to as the "spacer", and contains a motif with homology to the CDK inhibitors p21, p27, and p57. Indeed, both p107 and p130 are able to stably interact with cyclin A/CDK2 or cyclin E/CDK2 and to inhibit their kinase activity (Woo et al., 1997; Castano et al., 1998). Like the overexpression of p21 or p27, which arrests cells in G1 (Xiong et al., 1993; Toyoshima and Hunter, 1994), overexpression of p107 in some cells induces a G1 arrest, and this arrest is dependent on the integrity of the cyclin/CDK binding regions of p107.

Neither p107 nor p130 is frequently mutated in human tumors

In contrast to the frequent mutations of RB in human tumors, the evidence potentially implicating p107 or p130 in tumorigenesis is sparse. Chromosomal alterations in the p107 gene have been described in one cell line derived from a T-cell leukemia, one line derived from a B-cell lymphoma, and one primary tumor from an acute lymphoblastic leukemia (Takimoto et al., 1998). Full-length p107 protein was no longer detectable in either cell line. One point mutation at the
p130 locus has been described in a cell line derived from a small cell lung cancer (Helin et al., 1997). The possibility remains that the cell line mutations could have arisen during tissue culture; furthermore, the lack of tumor-free controls from the affected patients allow the possibility that the mutations may be unrelated to the genesis of the cancer. It is also possible that the inactivation of p107 or p130 may indeed contribute to tumorigenesis, if infrequently. Mutation of p107 in the mouse has variable consequences depending on the genetic background (Lee et al., 1996; LeCouter et al., 1998a). In no background has the loss of p107 been reported to promote tumorigenesis, although loss of p107 does induce hyperplasia in certain myeloid tissues of one mouse strain (LeCouter et al., 1998a). Chimeric mice with cells deficient in both pRB and p107 develop retinoblastoma at high frequency, suggesting that p107 may share a tumor-suppressive function with pRB, at least in this system (Lee et al., 1996). As seen for p107 mutations, mutations inactivating p130 have dramatically differing consequences in different genetic backgrounds, but these consequences do not seem to include an increased frequency of tumorigenesis (Cobrinik et al., 1996; LeCouter et al., 1998b).

*p107 and p130 in apoptosis, transformation, and cell cycle progression*

p107 and p130 associate with the same domains of E1A required for interaction of the latter with pRB (Dyson et al., 1992), raising the possibility that they might participate in the regulation of apoptosis, transformation, or the G1/S transition. Although mice lacking p107 are healthy and fertile, mice lacking both p107 and pRB demonstrate increased levels of apoptosis in the liver and CNS than mice only lacking pRB, indicating that they may share an apoptosis-suppressing function (Lee et al., 1996). Some mice in which the p130 gene has been mutated show extensive apoptosis in a variety of tissues, but this apoptosis is only observed in certain genetic backgrounds (LeCouter et al., 1998b).

In contrast, no direct evidence implicates either protein in the regulation of cellular transformation. p107 and p130, like pRB, do associate with other transforming viral oncoproteins such as SV40 large T antigen and HPV E7, and the inability of a T antigen mutant that has lost the
ability to bind pRB, p107 or p130 to transform wild-type or pRB-deficient MEFs may point to a transformation-suppressive function for p107 and/or p130. Interestingly, an additional domain of T antigen that promotes the degradation of p130 and changes in the phosphorylation state of p107 and p130 is also required to induce MEFs to grow to high density or in low serum; this domain is dispensable in MEFs deficient in both p107 and p130 (Zalvide et al., 1998).

The nature of the role played by these proteins in the regulation of the G1/S transition remains somewhat controversial. p130 is expressed primarily in G0 cells, although it is still detectable in G1 and S phase (Cobrinik et al., 1993). The expression of p107, in contrast, is strongly induced upon reentry of cells into the cell cycle (Moberg et al., 1996). Like pRB, p107 and p130 can be phosphorylated by G1-specific CDK complexes (Li et al., 1993; Beijersbergen et al., 1995). As mentioned above, overexpression of p107 can induce a G1 arrest. The effects of p107 or p130 loss on cellular proliferation in mice are strongly dependent on the genetic strain of the mice. In one genetic background, both MEFs and peripheral lymphocytes from these mice retain restriction point controls, although lymphocytes from mice lacking both p107 and p130 are hypersensitive to stimulation by concanavalin A and MEFs deficient in both proteins have been reported to enter S phase slightly earlier than wild-type cells following release from serum starvation (Cobrinik et al., 1996; Mulligan et al., 1998). In another background, loss of p107 leads to a doubling of the proliferation rate of MEF cultures, and loss of p130 in the mouse leads to extensive aberrant S phase entry in neural tissues. There exist, therefore, biological contexts in which p107 and p130 can play critical roles in proliferation control, although the genetic modifier or modifiers that must also contribute to proper cellular regulation remain unknown.

**C7. pRB suppresses growth through its “large pocket”**

pRB, p107, and p130 bind to viral oncoproteins through their conserved "small pocket" (Hu et al., 1990; Huang et al., 1990; Kaelin et al., 1990; Ewen et al., 1991; Hannon et al., 1993). Mutations in this region of pRB that prevent its interaction with viral oncoproteins also inhibit the ability of ectopically expressed pRB to suppress the growth of RB-deficient tumor-derived cell
lines (Qian et al., 1992; Qin et al., 1992; Hiebert, 1993). Similarly, although overexpression of wild-type pRB can inhibit apoptosis in SAOS-2 cells, expression of a mutant pRB that is unable to bind E1A is also unable to inhibit apoptosis in these cells (Haas-Kogan et al., 1995). This suggests a model in which E1A interferes with a function of pRB through direct binding to the domain of the protein required for that function.

While the small pocket is necessary for pRB-mediated growth suppression, it is not sufficient (Qian et al., 1992; Qin et al., 1992; Hiebert, 1993). The minimal region of pRB required to suppress growth includes additional sequences immediately to the carboxy-terminus of the small pocket (Qin et al., 1992; Hiebert, 1993); together with the small pocket, this region is referred to as the "large pocket", and is somewhat conserved with p107 and p130. All known tumor-derived mutations of RB disrupt the large pocket through deletion, truncation or point mutation (Sellers and Kaelin, 1996). This region, then, appears central to the function of pRB both in vitro and in vivo.

Like E1A, pRB is believed to carry out its biological activities through interactions with other proteins. Phosphorylation of the protein by CDK complexes, for example, seems to regulate the inactivation of the protein as the cell prepares to traverse the G1/S transition. While this reveals much about the proper regulation of the protein and the importance of the timing of its biological activity, it does not reveal how pRB exerts its activity on the cell. Although many proteins have been reported to bind to pRB in vitro, one particular cellular factor has been shown to stably interact with pRB both in vivo and in vitro in a manner specifically dependent on an intact, underphosphorylated large pocket: the transcription factor E2F.
D. E2F function is antagonized by pRB, p107, and p130

*pRB binds to E2F and inhibits cell proliferation*

The E2F transcription factor was originally identified as an E1A-responsive cellular transcription factor with a sequence-specific DNA binding activity (Kovesdi et al., 1986b; Kovesdi et al., 1986a; Kovesdi et al., 1987; Yee et al., 1987). In the context of adenoviral infection, E2F binds to two sites in the promoter of the viral E2 gene; these sites are essential for efficient expression of the gene (Kovesdi et al., 1986b). Not surprisingly, E2F sites have also been identified in a number of endogenous cellular genes, and have been shown to mediate the upregulation of these genes at the G1/S transition. Most of these genes are regulatory genes required in G1 or S, such as RB (Shan et al., 1994) or B-myb (Lam and Watson, 1993; Lam et al., 1994; Lam et al., 1995), or genes required for the machinery of DNA synthesis, such as dihydrofolate reductase (Means et al., 1992; Wade et al., 1992; Slansky et al., 1993; Wade et al., 1995) or DNA polymerase alpha (Pearson et al., 1991).

In cycling cells, a significant percentage of E2F is associated with pRB (Bandara and La Thangue, 1991; Chellappan et al., 1991). This interaction is disrupted by wild-type E1A, but not by mutant forms that have lost the ability to bind pRB. Furthermore, all known tumor-derived mutations of pRB disrupt its ability to bind E2F (Sellers and Kaelin, 1996). The inactivation of pRB function, either in tumor-derived mutations or in association with viral oncogenes, frees E2F from its interactions with pRB.

In normal cells, the association between pRB and E2F is regulated during the cell cycle. Hypophosphorylated pRB, which predominates during cell cycle arrest, can be found in association with E2F. pRB that has been hyperphosphorylated by cyclin/CDK complexes following passage through the restriction point (DeCaprio et al., 1989), in contrast, has lost its affinity for E2F (Chellappan et al., 1991; Knudsen and Wang, 1997). Mutations of the phosphorylation sites in pRB enhance the ability of the protein to induce a cell cycle arrest when overexpressed, but this arrest can be overcome by the overexpression of an E2F subunit (Qin et
Thus, hypophosphorylated pRB binds E2F and inhibits cell proliferation, whereas the hyperphosphorylation of pRB leads to the dissociation of E2F and passage through the G1/S transition.

D1. The E2F and DP gene families

Although E2F was originally described as a DNA-binding activity, it is in fact rather heterogeneous, resulting from the proteins encoded by at least seven distinct genes in mammals (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992; Girling et al., 1993; Ivey-Hoyle et al., 1993; Lees et al., 1993; Beijersbergen et al., 1994; Ginsberg et al., 1994; Hjimans et al., 1995; Ormondroyd et al., 1995; Sardet et al., 1995; Wu et al., 1995; Zhang and Chellappan, 1995; Rogers et al., 1996). These are grouped into two families of genes, referred to as E2F (E2F-1 through -5) and DP (DP-1 and -2). Either of the DP polypeptides are capable of heterodimerizing with any of the E2F polypeptides; each of the possible ten resulting pairs are observed in cell extracts. Heterodimerization is essential for efficient binding to DNA and transcriptional activation by the complex (Bandara et al., 1993; Helin et al., 1993; Krek et al., 1993).

Whereas the sequence conservation between the two DP genes is high throughout the length of their coding sequences, the sequence homology among the E2F genes is concentrated in a few key domains. The domain encoding the DNA-binding function is the most highly conserved, and does not resemble DNA binding domains from other transcription factors. This domain is followed immediately by a domain required for efficient heterodimerization with the DP proteins. The transcriptional activation function of the complex is encoded by a conserved region in the 3′ end of the E2F genes. Based on sequence conservation and gene structure, E2F-4 and -5 form a subfamily of E2F proteins distinct from the subfamily including E2F-1, -2, and -3. This latter group contains an amino-terminal region that includes a conserved nuclear localization signal (Verona et al., 1997); this region is absent in E2F-4 and -5.

Domains required for DNA binding and dimerization are also found in the DP proteins, and these regions show limited homology to the corresponding regions of the E2F proteins.
domain in the DP proteins contributing to transcriptional activation by the complex (other than the dimerization and DNA-binding domains) has been identified; rather, presence of the DP moiety in the E2F/DP complex is essential for DNA binding and the potentiation of the transcriptional activation function of E2F. The DP polypeptide also promotes the interaction of the complex with pRB, although this binding occurs primarily through the transcriptional activation domain of the E2F subunit.

D2. pRB-mediated repression

The association of pRB with E2F is regulated during the cell cycle. This regulation is believed to occur both through the specific phosphorylation of pRB at the G1/S transition by CDK complexes and through transcriptional regulation of the RB gene. pRB/E2F complexes cannot be detected in quiescent cells, but can be detected in late G1 as these cells reenter the cell cycle and the transcription of pRB is activated (Ikeda et al., 1996; Moberg et al., 1996). They can also be detected in S phase. This is believed to result from newly synthesized pRB, not yet hyperphosphorylated by CDK complexes, binding to E2F. Importantly, when cells are induced to arrest and differentiate, pRB is no longer present in a primarily hyperphosphorylated form, and a much greater percentage of E2F complexes associate with pRB (Ikeda et al., 1996).

When bound to E2F, pRB inhibits E2F-mediated transcription. This was first indicated by studies of the E2 promoter: overexpression of pRB inhibits the expression of the promoter in a manner dependent on the presence of the E2F sites, and this effect is only observed with alleles of pRB that are competent to bind E2F (Hiebert et al., 1992). The repression of the E2 promoter can be alleviated by co-expression of E1A or HPV E7 (Hiebert et al., 1992). This derepression can be explained by the in vitro and in vivo observations that E1A induces the dissociation of pRB from E2F complexes (Bagchi et al., 1990). pRB does not prevent DNA binding by the complex; endogenous pRB-E2F complexes are easily detectable in gel shift assays (Chellappan et al., 1991; Moberg et al., 1996). The inhibitory activity of pRB is instead specifically a repression of transcription by the intact DNA-bound complex.
**pRB represses basal transcription**

The presence of pRB, however, does not merely mask the transcriptional activation domain of the E2F moiety, but also represses the basal level of transcription when present at a promoter. This was first suggested by the observation that the overexpression of pRB inhibits the E2 promoter more than the simple deletion of the E2F sites does, and this additional inhibition is dependent on the integrity of the E2F sites (Hiebert et al., 1992). Similarly, it has been shown that the insertion of E2F sites can repress artificial promoters in cells with wild-type pRB, whereas the same insertion activates the expression of those promoters in cells without functional pRB, or in cells expressing wild-type E1A (Weintraub et al., 1992). That this repression is the direct result of the action of pRB on the promoter has been demonstrated with fusion proteins between pRB and the DNA-binding domain of the yeast GAL4 protein: this fusion can repress the expression of artificial promoters in a manner dependent on the presence of either a GAL4 site (to which GAL4-pRB could directly bind) or an E2F site (to which GAL4-pRB could bind through an E2F intermediary) (Adnane et al., 1995; Bremner et al., 1995; Weintraub et al., 1995). Thus, although "free E2F", E2F/DP complexes not associated with pRB, can activate transcription, pRB-E2F complexes repress basal levels of transcription.

The mechanisms of pRB-mediated repression are under active investigation. Repression requires only the "small pocket" of pRB, and can be alleviated by overexpression of viral oncoproteins or of cyclins A or E. Repression may be effected by interactions between pRB and other promoter elements (Dyson, 1998) and/or by recruitment of histone deacetylase activity to the promoter by pRB (Brehm et al., 1998; Luo et al., 1998; Magnaghi-Jaulin et al., 1998).
D3. p107 and p130 also bind E2F and can repress transcription

pRB, p107, and p130 bind E2F at distinct times during the cell cycle

Although "free E2F" and pRB-associated E2F account for a significant percentage of the E2F in the cell, other E2F complexes exist as well. Like the pRB association, these complexes can be disrupted by E1A (Bagchi et al., 1990). The cloning of the E1A-associated proteins p107 and p130 and their significant homology to pRB led rapidly to their identification as components of these "other" E2F complexes (Ewen et al., 1991; Cao et al., 1992; Shirodkar et al., 1992; Cobrinik et al., 1993).

E2F complexes containing p107 and those containing p130 show distinctly different patterns of appearance and disappearance during the cell cycle. p130/ E2F complexes are present at their highest levels in quiescent cells, and become increasingly less abundant as cells reenter the cell cycle, although they can still be detected as cells enter S phase (Cobrinik et al., 1993; Ikeda et al., 1996; Moberg et al., 1996). p107/ E2F complexes, in contrast, are not detectable in quiescent cells, first appearing in late G1 and reaching their highest levels in S phase (Moberg et al., 1996). Binding of p107 or p130 to CDK complexes and E2F is not mutually exclusive; the five polypeptide complex p107 or p130/ cyclin A or E/ CDK 2/ E2F/ DP complex can be observed bound to DNA in a gel shift assay, although the potential biological function of a presumably inactive kinase in this complex remains unknown. Like pRB, p107 and p130 can repress E2F activity (Bandara et al., 1993; Zamanian and La Thangue, 1993; Bremner et al., 1995; Ferreira et al., 1998; Pierce et al., 1998b) and may be associated with a histone deacetylase (Ferreira et al., 1998). p130 is therefore proposed to mediate the repression of E2F target genes in quiescent cells; the role p107/ E2F complexes might be playing, at times in the cell cycle when known E2F target genes are actively expressed, remains unclear.
pRB, p107, and p130 bind to distinct forms of E2F

It has been proposed that the failure to detect specific roles for p107 or p130 in apoptosis or growth control in some contexts could result from extensive redundancy of function among pRB, p107, and p130. In this model, the three proteins would share a common suppressive function, such as transcriptional repression at E2F DNA-binding sites, and only the inactivation of all three genes could reveal their importance to the biology of the cell. Alternatively, the differences in the biology of pRB, p107, and p130 could result from their associations with biochemically distinct forms of E2F. Indeed, pRB can be found in association with E2F-1, -2, -3, or -4 (Moberg et al., 1996), whereas p107 is associated only with E2F-4 (Moberg et al., 1996) and p130 with both E2F-4 and E2F-5 (Hijmans et al., 1995; Sardet et al., 1995); although the DP subunit is always present in these complexes, the identity of the subunit, DP-1 or DP-2, does not seem to affect the specificity of the interaction.

A different subset of E2F transcriptional targets is deregulated in MEFs lacking p107 and p130 than the subset deregulated in pRB-deficient MEFs. Cyclin E expression, for example, is upregulated in pRB-deficient MEFs, but not in p107/p130-deficient cells (Herrera et al., 1996; Hurford et al., 1997). B-myb, in contrast, is derepressed in p107/p130-deficient MEFs, but shows a wild-type pattern of expression in pRB-deficient cells (Hurford et al., 1997). These proteins therefore regulate different biochemical pathways in these cells. Given that pRB, p107 and p130 associate with different E2F proteins, this specific deregulation could result from differential promoter selection mediated either by the E2F subunit or by interactions between the pRB homolog and other promoter elements.
E. **E2F promotes proliferation and apoptosis**

E1. **E2F promotes progression through G1**

Like E1A, ectopic expression of the E2F genes promotes entry into S phase. E2F-1, -2, -3, -4, and -5 all have this ability in some experiments, although neither E2F-4 nor -5 efficiently promotes S phase in others (Johnson et al., 1993; Beijersbergen et al., 1994; Lukas et al., 1996). Overexpression of E2F-1 can override or prevent G1 arrest induced by a variety of stimuli, including transforming growth factor-beta, gamma-irradiation, and the overexpression of CDK inhibitors p16, p21, p27 or dominant-negative CDK2 (DeGregori et al., 1995; Johnson et al., 1993; Lukas et al., 1996; Mann and Jones, 1996; Schwarz et al., 1995). Overexpression of E2F-1 in mouse megakaryocytes also induces their hyperproliferation and interferes with their terminal differentiation (Guy et al., 1996). S phase promotion by E2F-1 requires both its DNA-binding and transactivation functions (Johnson et al., 1993; Qin et al., 1994), whereas mutations in E2F-1 with reduced affinity for pRB enhance the ability of E2F-1 to promote S phase entry (Shan et al., 1996).

In addition to their ability to induce inappropriate entry into S phase when misexpressed, transcriptional activation via the E2F genes may also be essential for the G1/S transition. Overexpression of dominant-negative forms of DP-1 or DP-2 can promote the accumulation of SAOS-2 cells in G1 (Wu et al., 1996; Fan and Bertino, 1997). Similarly, microinjection of a dominant-negative form of E2F-1 can prevent the stimulation of S phase entry by E1A (Dobrowolski et al., 1994), and inhibition of E2F DNA-binding activity inhibits S phase induction following serum stimulation (Ishizaki et al., 1996). Although all five E2F genes share the ability to induce S phase entry when overexpressed, microinjection of an antibody to E2F-3 inhibits S phase progression in a rat embryo fibroblast cell line (REF52), whereas an antibody to E2F-1 was without effect, suggesting an specific role for E2F-3 in G1/S progression (Leone et al., 1998). E2F-3-deficient MEFs proliferate more slowly than wild-type MEFs; however, they do proliferate,
and mice that lack E2F-3 are fully viable, arguing that E2F-3 is rate-limiting for progression but not irreplaceable (P. Humbert, K. Rogers, S. Dandapani, and J. Lees, unpublished data).

**E2F complexes promote cell cycle progression in Drosophila**

Homologs of E2F, DP, and pRB have also been identified in *D. melanogaster* (Dynlacht et al., 1994a; Ohtani and Nevins, 1994; Du et al., 1996a; Sawado et al., 1998), where they also seem to be involved in the regulation of cell proliferation and S phase progression. Loss-of-function mutations in an E2F homolog, *dE2F*, or a DP homolog, *dDP*, are lethal, usually at the late larval or pupal stages of development, and are associated with a slower progression through S phase (Royzman et al., 1997; Duronio et al., 1998). Elimination of *dE2F* function in a compartment of the developing wing has also been shown to slow cell cycle progression, as has overexpression of the RB homolog *RBF* (Neufeld et al., 1998). In contrast, the ectopic expression of *dE2F* induces an unscheduled entry into S phase in a variety of *Drosophila* tissues (Asano et al., 1996; Duronio et al., 1996), and in the wing promotes rapid cellular proliferation; this effect is further enhanced by coexpression of *dDP* (Neufeld et al., 1998). The promotion of cellular proliferation by E2F and DP proteins is therefore a function highly conserved across metazoan evolution.

Thus, *Drosophila* embryos and larvae do not require zygotic *dE2F* or *dDP* expression for progression, albeit slow, through S phase. If the maternal stores of E2F-encoding mRNA have been exhausted, this suggests that some components of E2F activity are partially dispensable for cell cycle progression in this organism during its early development. A second *Drosophila* E2F homolog, termed *dE2F2*, has recently been described (Sawado et al., 1998), and the potential existence of additional E2F or DP homologs cannot yet be dismissed; the mutations in *dE2F* and *dDP* are therefore not necessarily eliminating all E2F complexes in these cells. Furthermore, additional arguments suggest that the expression of transcriptional targets of E2F may indeed be required for S phase progression in *Drosophila*. First, some *Drosophila* target genes known to be regulated by *dE2F*, such as cyclin E and PCNA, are also known to be required for S phase (Henderson et al., 1994; Knoblich et al., 1994). Second, some E2F transcriptional targets are
derepressed by loss-of-function mutations in dDP (Royzman et al., 1997; Duronio et al., 1998), raising the possibility that, although E2F-mediated transcriptional activation is clearly requisite for efficient progression through S phase, derepression of E2F target genes may suffice for passage through S. Thus, although dE2F and dDP are required for efficient passage through S phase, the question of a general requirement for the induction of E2F transcriptional targets in Drosophila S phase progression remains unresolved.

E2. E2F promotes transformation and apoptosis

E2F genes also have some capacity to promote cellular transformation in vitro. Overexpression of E2F-1 has been reported to cooperate in a transformation assay with a mutant form of HPV E7 that has lost its ability to bind to the pRB family, suggesting that activation of the E2F pathway may be the critical role of this domain in transforming viral oncoproteins (Melillo et al., 1994). Overexpression of the DP proteins, either alone or in combination with E2F-1 or E2F-4, has been shown to cooperate with activated ras in transformation assays. (Beijersbergen et al., 1994; Johnson et al., 1994a; Jooss et al., 1995), although they seem to cooperate less well with ras than c-myc or E1A does in these assays. Overexpression of E2F-1, -2, or -3 in NIH 3T3 cells also promotes transformation (Xu et al., 1995). Mutational studies of E2F-1 indicate that its transforming activity requires an intact DNA-binding domain (Xu et al., 1995), although a heterologous transcriptional activation domain can substitute for the endogenous one (Johnson et al., 1994a).

E2F-1 promotes the induction of apoptosis

The activation of E2F-1 also promotes apoptosis. Overexpression of E2F-1 induces apoptosis in a variety of systems (Qin et al., 1994; Shan and Lee, 1994; Wu and Levine, 1994; Kowalik et al., 1995). Ectopic expression of the Drosophila protein dE2F also induces apoptosis in multiple cell types (Asano et al., 1996; Du et al., 1996b; Neufeld et al., 1998), indicating that, like the promotion of cellular proliferation, this function has been well conserved during evolution.
Much of the apoptosis induced by the overexpression of E2F-1 in mammalian systems depends on the presence of functional p53 protein (Qin et al., 1994; Wu and Levine, 1994; Kowalik et al., 1995). The observation that ectopic expression of E2F-1 induces expression of p19-ARF (DeGregori et al., 1997; Bates et al., 1998), a protein that activates p53-mediated transcription and stabilizes p53 protein levels (Pomerantz et al., 1998; Zhang et al., 1998) suggests that at least some of the effects of E2F-1 on apoptosis may be mediated by this pathway. E2F-1 overexpression has also been shown to promote apoptosis in some systems lacking functional p53 (Hsieh et al., 1997; Phillips et al., 1997). The effects of E2F-1 on p53-mediated apoptosis depend on the DNA-binding and transcriptional activation functions of E2F-1, although p53-independent apoptosis can be induced in SAOS-2 cells by a form of E2F-1 lacking a transcriptional activation domain.

Ectopic expression of E2F-1 induces p53-dependent apoptosis in keratinocytes of transgenic mice (Pierce et al., 1998a), and p53-independent apoptosis in the testes (Holmberg et al., 1998).

The induction of apoptosis is not a general consequence of E2F activation. Rather, the ability to promote apoptosis is limited to particular E2F complexes. Ectopic expression of E2F-1 in REF52 cells induces apoptosis, whereas expression of E2F-2, -3, -4, or -5 does not (DeGregori et al., 1997). Similarly, apoptosis is induced in an astrocytoma cell line by ectopic expression of either E2F-1 or E2F-2, but not by the other E2F genes (Dirks et al., 1998). In both of these cases E2F overexpression, sometimes in combination with DP-1, was observed to promote cell cycle progression regardless of which E2F was used in the experiment; only the ability to induce apoptosis appears specific in these assays.

Experiments in mice indicate that E2F-1 may regulate apoptosis in a variety of tissues. Mice lacking functional E2F-1 develop thymic hypertrophy, and thymocytes from these mice are partially resistant to apoptosis (Field et al., 1996). Homozygous deletion of the E2F-1 gene also rescues most of the aberrant apoptosis and S phase entry observed in the lens and CNS of mice lacking pRB, although aberrant apoptosis in the PNS is still observed (Tsai et al., 1998). Similarly, loss of E2F-1 rescues apoptosis in transgenic mice expressing in the choroid plexus a fragment of SV40 large T antigen (T-121) retaining its ability to bind pRB-family members (Pan et
Thus, E2F-1 activity promotes the induction of apoptosis *in vivo* as well as *in vitro*. Once again, this suggests some specificity of E2F function.

**E3. E2F as an oncogene, and as a tumor suppressor**

Although the pRB pathway is routinely mutated in human cancers and although the expression of cellular or viral oncogenes can lead to an activation of E2F transcriptional activity, the E2F genes themselves do not seem to be mutational targets in human tumors. Alterations in the length of a polyserine repeat in E2F-4 have been reported, although this was believed to result from the genomic stability associated with tumorigenesis and not believed to have been a causative event in tumorigenesis (Yoshitaka et al., 1996; Souza et al., 1997; Walsh et al., 1998). A variety of hypotheses may explain the lack of tumor-promoting E2F mutations in tumors. The transcriptional targets of multiple E2F genes may promote tumorigenesis, in which case mutations in regulators of E2F, such as pRB, may be strongly favored statistically, because these mutations simultaneously deregulate the products of multiple E2F genes. Another possibility is that mutations in E2F genes that abolish binding to pRB-family members without reducing transcriptional activity may be rare or even impossible (R. Verona and J. Lees, unpublished data).

Experiments with E2F-1 deficient mice, however, do indicate a role for this gene in tumorigenesis. E2F-1 deficiency prevents thyroid tumorigenesis in mice heterozygous for an RB null mutation, and reduces the incidence of pituitary tumorigenesis (Yamasaki et al., 1998). E2F-1 deficiency also inhibits cellular proliferation in T-121 induced tumors of the choroid plexus (Pan et al., 1998). A function supplied by E2F-1 is therefore essential for efficient tumorigenesis in some contexts, indicating that the deregulation of E2F by mutagenesis of the pRB/ cyclin D/ CDK4/ p16 pathway is highly relevant to tumorigenesis affected by this pathway (Figure 1). Mice deficient for E2F-1 develop late onset tumors of the reproductive tract more often than wild-type mice, indicating that E2F-1 can also act as a tumor suppressor (Yamasaki et al., 1996). It has not yet been determined if this suppressive role for E2F-1 is cell autonomous. If it is, it could indicate that removal of a pRB-E2F-1 complex derepresses an oncogenic target gene; alternatively,
transcriptional activation by E2F-1 might be necessary for the expression of a tumor suppressor in these tissues, perhaps related to the pro-apoptotic function of E2F-1.
F. E2F genes: specificity in function and in regulation

F1. Individual E2F genes have specific functions

The studies of E1A have led to a transcriptional activity, E2F, which participates in the regulation of cellular proliferation and apoptosis both in normal tissues and in tumors. To understand the mechanisms through which E2F regulates these processes requires the identification of the specific E2F genes involved in the regulation of each process; it also requires an understanding of the ways specific E2F proteins are regulated.

The existence of some functional overlap among E2F genes seems very likely; for example, all five E2F genes have been reported to promote cell cycle progression when overexpressed, although in some cases this effect requires the coexpression of a DP subunit. There is, nevertheless, strong evidence that individual E2F genes also play distinct roles in the regulation of cell biology. As discussed previously, E2F-1 is involved in the regulation of apoptosis, a role not indicated for the other E2F genes with the possible exception of E2F-2. E2F-1 may also play a specific role in the deregulation of the cell cycle caused by loss of pRB. E2F-3 may also have a specific role in G1/S progression, as indicated by the proliferative defect in E2F-3-deficient MEFs and the ability of an E2F-3-specific antibody to inhibit S phase entry. Unlike mice deficient for E2F-1 or -3, mice without E2F-4 display hematopoietic defects (P. Humbert, K. Rogers, and J. Lees, unpublished data). Mice without E2F-5 die at about six weeks of age with hydrocephalus, apparently the result of excessive cerebrospinal fluid production in the choroid plexus (Lindeman et al., 1998). The defects observed in mice lacking E2F-4 or E2F-5 may suggest a role for E2F complexes in differentiation, or may be consequences of cell cycle deregulation in certain tissue types. Regardless of the mechanisms involved, specific E2F genes must supply a variety of non-redundant functions important at the cellular and organismal levels. This is consistent with the argument that pRB, p107, and p130 differ in their biological activities because they regulate different subsets of E2F proteins (Figure 2). pRB may suppress tumorigenesis because the E2F proteins regulated by pRB may have specific functions relevant to tumorigenesis, whereas neither
Figure 2. The pRB family members inhibit different E2F proteins with distinct functions.

Inhibitory interactions are depicted by the lines connecting pRB and E2F family members.

Functions of each protein are listed below the name of the protein.
p107 nor p130 is frequently lost in cancer because the E2F proteins they regulate may have functions less likely to promote tumorigenesis. While this hypothesis remains speculative, its underlying premise that the E2F proteins differ in some of their functions is apparently true.

**F2. The E2F genes are regulated in overlapping but distinct ways**

Given that the E2F proteins have non-overlapping functions in addition to any shared functions, it is perhaps unsurprising that there are differences in their methods of regulation. Obviously, the best characterized method of regulation is the association of E2F complexes with pRB-family members: E2F-5 associating only with p130, E2F-1, -2, and -3 associating only with pRB, and E2F-4 associating with pRB, p107, and p130. The observation that the three pRB-family members have both distinct and overlapping functions is therefore consistent with the distinct and (probably) overlapping functions observed for the E2F proteins they regulate. Mutation of the pRB-family members deregulates specific subsets of E2F target genes; these subsets, then, presumably mediate the specific biological effects of E2F and pRB family members on cell biology.

_E2F mRNA levels are regulated during development_

Individual E2F family members are also regulated by control of their expression, localization, and DNA binding activities. E2F-2 and E2F-4, for example, are widely expressed in the developing mouse (Dagnino et al., 1997a). All five E2F transcripts can be detected by in situ hybridization in the kidney, liver, lung, and heart of the developing mouse (Dagnino et al., 1997a). E2F-5, however, appears absent from the developing thymus and from epithelial tissues that are not undergoing terminal differentiation (Dagnino et al., 1997a; Dagnino et al., 1997b). E2F-3 is also undetectable in the developing thymus, and neither E2F-1 nor E2F-3 can be detected in the developing choroid plexus or in chondrocytes (Dagnino et al., 1997a). In skeletal muscle, only E2F-2 can be readily detected by in situ hybridization, although E2F-3, -4, and -5, but not E2F-1, can also be detected (Dagnino et al., 1997a). Whether this control over E2F mRNA levels during...
development results from transcriptional control or from regulation of mRNA stability remains unresolved. Regardless of the mechanism(s) involved, the restricted expression of particular subsets of E2F genes in some tissues is intriguing and would be consistent with the inactivation of certain E2F-dependent regulatory pathways in a tissue-specific manner.

**E2F levels and localization are regulated during the cell cycle**

Much more is known about the regulation of E2F proteins from studies of cells grown in vitro. The primary E2F complexes in quiescent cells are E2F-4/ DP and E2F-5/ DP, each of which are associated with p130 and therefore assumed to be repressing the expression of E2F target genes; the E2F-4-containing complexes outnumber the E2F-5 complexes. As cells enter the cell cycle and approach S phase, E2F-4 and E2F-5 become increasingly excluded from the nucleus, such that by S phase the vast majority of E2F-4 and -5 complexes are cytoplasmic (Lindeman et al., 1997; Muller et al., 1997; Verona et al., 1997). This relocation of E2F-4 may promote the derepression of E2F target genes at the approach of the G1/S transition, as E2F-4 is primarily associated with p130 as cells are first reentering the cell cycle and with pRB and p107 in late G1 and S (Moberg et al., 1996). The relocation of both proteins may also prevent the premature activation of E2F target genes in early- to mid-G1 when free E2F-4 complexes first appear. E2F-4 and -5 are the only E2F proteins known to undergo changes in subcellular localization; E2F-1, -2, and -3 all seem to be constitutively nuclear (Verona et al., 1997).

Reentry of cells into the cell cycle is also associated with the de novo synthesis of E2F proteins. Whereas E2F-4 levels are only mildly elevated upon exit of cells from a quiescent state, levels of E2F-1, -2, and -3 are strongly induced (Hsiao et al., 1994; Johnson et al., 1994b; Neuman et al., 1994; Moberg et al., 1996). The induction of E2F-1 and -2 appears to occur at the transcriptional level and are mediated, at least in part, by the presence of E2F binding sites in their promoters (Hsiao et al., 1994; Johnson et al., 1994b; Neuman et al., 1994; Sears et al., 1997); E2F-3 is also upregulated at the transcriptional level, although this upregulation is much less dramatic than the induction of its protein levels or DNA binding activity (Moberg et al., 1996; Leone et al., 1998). This discrepancy may be explained by the observation that the E2F-3 protein
in NIH 3T3 cells that have been blocked in S phase is 40-fold more stable than in asynchronous cells (Flores et al., 1998). Although E2F-1 and E2F-2 expression and DNA binding activities are strongly induced upon reentry of cells into the cell cycle, accumulation of E2F complexes at the G1/S transition in continuously cycling cells has only been described for E2F-3.

Protein stability may also regulate other members of the E2F family. At least E2F-1 and E2F-4, when overexpressed, are targeted for degradation by the ubiquitin-proteasome pathway (Hateboer et al., 1996; Hofmann et al., 1996). Coexpression of pRB-family members competent to bind the E2F in question results in the stabilization of the E2F complex; this is postulated to result from the masking of a carboxy-terminal epitope on the E2F moiety required for its targeting and/or degradation. If endogenous E2F complexes are regulated in the same manner, this would suggest that the levels of free E2F complexes are reduced by their specific degradation, whereas complexed E2F is relatively stable in comparison. Whether or not the degradation of free E2F complexes is regulated in a cell cycle-dependent manner, beyond the cell cycle-dependent pRB-family associations, is unknown.

Some E2F complexes are inactivated by specific phosphorylation

The DNA binding activity of some E2F complexes is regulated by phosphorylation. E2F-1 can interact with cyclin A through the same region of E2F-1 required for its nuclear localization (Dynlacht et al., 1994b; Krek et al., 1994; Xu et al., 1994; Kitagawa et al., 1995; Krek et al., 1995; Dynlacht et al., 1997). This kinase is active, and phosphorylates in vitro both subunits of an E2F-1/DP-1 complex on sites that are also phosphorylated in vivo. Cyclin A-mediated phosphorylation of the DP-1 subunit in this complex leads to a loss of DNA binding activity by the complex (Dynlacht et al., 1994b; Krek et al., 1994; Xu et al., 1994; Kitagawa et al., 1995). E2F-2 and E2F-3 are also associated with a kinase activity in vivo (Krek et al., 1994), and incubation of E2F-3/DP-1 complexes with cyclin A/CDK2 inhibits their DNA binding activity (Dynlacht et al., 1997), suggesting that this downregulation may be shared amongst these three E2F complexes. E2F-4/DP-1 and E2F-5/DP-1 complexes, in contrast, are unaffected by treatment with cyclin A/
CDK2. Downregulation of E2F-1/DP-1 complexes by cyclin A/CDK2 requires the cyclin A binding domain of E2F-1; transplantation of this domain to the amino-terminus of E2F-4 confers to E2F-4/DP-1 complexes susceptibility to cyclin A/CDK2-mediated inhibition (Dynlacht et al., 1997). Interestingly, although CDK complexes containing cyclin B are capable of phosphorylating E2F-1/DP-1 complexes on the same residues as cyclin A complexes do, cyclin B cannot bind stably to E2F-1 and cannot inhibit DNA binding by E2F-1/DP-1 complexes (Dynlacht et al., 1997), indicating that phosphorylation of DP-1, although necessary, is not sufficient to inhibit DNA binding activity.

The activation of E2F target gene expression at the approach of the G1/S transition, then, is mediated by phosphorylation and release of the pRB-family members and de novo synthesis of E2F-1, -2, and -3. The inactivation of E2F target gene expression in S phase is effected by the relocalization of E2F-4 and -5 to the cytoplasm and the inhibition of E2F-1, -2, and -3 by cyclin A/CDK2. The importance of this inactivation is underlined by the observation that overexpression of a form of E2F-1 lacking the cyclin A-binding domain or a form of DP-1 which cannot be phosphorylated delays S phase progression (Krek et al., 1995). Only alleles of E2F-1 that retain DNA binding activity can induce this delay; alleles also retaining the transcriptional activation domain induce apoptosis following the S phase delay, inducing apoptosis much more efficiently than wild-type E2F-1 does. Inactivation of E2F activity may, therefore, be required for efficient progression through S phase and completion of the cell cycle.

Thus, the E2F genes are not redundant: they have individual functions in addition to any shared functions, and are regulated in overlapping but distinct ways. This underscores the importance of characterizing the functions of each member of the E2F gene family. In addition, antibody experiments indicate that E2F-1, -2, -3, -4, and -5 are insufficient to account for all of the E2F DNA binding activity in proliferating cells. The identification of the remaining members of the E2F family is obviously critical to any complete understanding of the functions of E2F in cell biology.
G. E2F genes: transcriptional targets

Given that E2F activity is rate-limiting for S phase entry under a variety of conditions, it is of great interest to identify the transcriptional targets of E2F promoting cellular proliferation. E2F has been proposed to regulate many genes expressed at the G1/S transition, such as pRB, p107, E2F-1, E2F-2, cyclin E, cyclin A, B-myb, DHFR, DNA polymerase alpha, HsOrc1, and cdc6. Both cyclin E and cyclin A are required for the G1/S transition, and cdc6 is required for DNA synthesis, as are, presumably, DHFR, DNA polymerase alpha, and perhaps HsOrc1. If all of these genes, as well as others which have not yet been identified, are legitimate and direct E2F targets, it may be very difficult to establish a target gene or genes that can substitute for E2F activity in the promotion of DNA synthesis.

Cyclin E

Cyclin E has, nevertheless, been intensively studied as a candidate for such a role. Cyclin E levels are induced in pRB-deficient MEFs (Herrera et al., 1996; Hurford et al., 1997) and in cells overexpressing E2F-1, -2, -3, or -4 (DeGregori et al., 1997), and cyclin E is rate-limiting for S phase progression in a wide variety of cells. Overexpression of cyclin E can overcome a G1 arrest imposed by overexpression of p16, by a form of pRB lacking phosphorylation sites, or by dominant-negative DP-1, and does so without a significant induction of E2F activity. Although the lack of E2F activity would probably lead to the eventual depletion of nucleotide pools for DNA synthesis and thereby prevent further cellular proliferation, overexpression of cyclin E in this system is apparently sufficient to permit at least one transition past the restriction point (Lukas et al., 1997). In contrast, studies on irradiated cells, in which elevated levels of CDK inhibitors lead to the inactivation of cyclin E-associated kinase activity and G1 arrest, have shown that overexpression of E2F-1 can permit entry into S phase without (full) rescue of cyclin E-associated kinase activity (DeGregori et al., 1995). In this system, then, hyperactivation of other E2F targets can substitute for activation of cyclin E. While these two findings are difficult to reconcile, it has been suggested that normally cycling cells may require both cyclin E and another E2F target for
timely progression through the restriction point, and that hyperactivation (perhaps unphysiological) of either can be sufficient for S phase entry in some circumstances (Lukas et al., 1997; Dyson, 1998). These results are reminiscent of data regarding cyclin E and cyclin D1: both are required in normally cycling cells in tissue culture, overexpression of either can accelerate S phase entry, and overexpression of both causes a further acceleration.

The relative contributions of E2F activity and cyclin E in the promotion of S phase entry have also been studied in D. melanogaster. As indicated above, mutations of dE2F or dDP retard S phase progression, whereas null mutations of cyclin E prevent almost all DNA synthesis following the first sixteen cell cycles in early Drosophila embryogenesis (Knoblich et al., 1994). Cyclin E mRNA levels are highly dependent on dE2F and dDP in the endocycling cells of the larval gut (Duronio et al., 1995; Royzman et al., 1997; Duronio et al., 1998). In contrast, the mRNA levels of E2F target genes in the developing Drosophila central nervous system depend on the expression of cyclin E (Duronio et al., 1995). Thus, as in mammals, cyclin E appears to be both an activator or E2F, perhaps through phosphorylation of the Drosophila RB homolog RBF, and a transcriptional target of E2F. In quiescent epidermal cells of the early Drosophila embryo, the induction of S phase by ectopic expression of dE2F and dDP is not observed in embryos lacking functional cyclin E, although the mRNA levels of E2F target genes are efficiently induced, indicating that the overexpression of dE2F with dDP cannot bypass the S phase requirement for cyclin E in this tissue (Duronio et al., 1996). Ectopic expression of cyclin E induces S phase in this tissue even in embryos with a mutation in dE2F (Duronio et al., 1995). This induction, however, is not reported to occur at a stage in embryogenesis at which the epidermal cells have been quiescent for a prolonged period of time, suggesting that dE2F transcriptional targets other than cyclin E may also be important in the promotion of S phase. Thus, in both Drosophila and mammals, cyclin E is a very important transcriptional target for E2F-mediated promotion of cell cycle progression, but does not seem to be the sole important target.
Another target proposed to participate in the promotion of S phase entry by E2F is cdc6. In yeast, cdc6 is directly involved in the initiation of DNA synthesis, and the same is believed to be true in humans. The human cdc6 mRNA is most abundant at the G1/S transition, at least in part as a result of cell cycle-dependent transcription (Leone et al., 1998). The variations in cdc6 transcription depend on two E2F binding sites in the cdc6 promoter (Hateboer et al., 1998; Ohtani et al., 1998; Yan et al., 1998). Both sites are occupied throughout the cell cycle, and their mutation causes constitutive, low levels of expression from the promoter, preventing both repression of the promoter in arrested cells and activation of the promoter in S phase cells. Microinjection of an antibody to cdc6 prevents entry into S phase, whereas the overexpression of cdc6 in cycling cells increases the percentage of cells in S phase, as does overexpression of cyclin E; coexpression of both increases this percentage even further. A more interesting question is whether or not cdc6 can induce S phase entry in quiescent cells, either alone or in cooperation with cyclin E, although there are some indications that cdc6 may be unstable in quiescent cells (Hateboer, 1998).

**G1. Some E2F targets are regulated by repression, some by activation**

The existence of non-overlapping functions for individual E2F genes and their differential regulation by transcription, localization, phosphorylation, stability and associated proteins raises the possibility that not all E2F target genes will be affected by E2F in the same way. Indeed, some evidence suggests this may be the case. For example, deletion of the E2F site in the mouse DHFR promoter abolishes activation of the promoter in S phase without affecting expression levels in early G1 (Means et al., 1992). In contrast, the E2F site in the B-myb promoter is required for repression in early G1, but not for peak expression in S phase (Lam and Watson, 1993) and the E2F sites in the cdc6 promoter may be required both for repression and for activation (Hateboer et al., 1998). The B-myb results are consistent with in vivo footprinting data indicating that its E2F site is occupied in quiescent fibroblasts, but that this occupation is no longer detectable once the
cells enter S phase (Zwicker et al., 1996), and with the derepression of B-myb observed in MEFs deficient for both p107 and p130. Thus, E2F sites in specific promoters do not all confer the same regulation to those promoters, although they do seem to share conferral of maximal promoter activity at the G1/S transition.

**G2. Individual E2F genes may regulate specific subsets of transcriptional targets**

Overexpression studies are also consistent with individual E2F complexes preferentially regulating certain promoters. For example, ectopic expression of E2F-1 activates the endogenous p19-ARF gene in REF52 cells, whereas E2F-2 activates it more weakly and the other E2F genes do not activate it (DeGregori et al., 1997). In contrast, E2F-2 overexpression is the most efficient at activating DHFR, E2F-3 is the most efficient activator of CDK2, and E2F-1, -2, -3, and -4 share the ability to activate cyclin E expression. In many overexpression assays, specificity of function can be lost; for example, E2F-2 will bind to p107 when overexpressed, although this interaction is not believed to occur in vivo (Moberg, 1998). That individual E2F genes could have demonstrably different effects on particular transcriptional targets even when overexpressed strongly suggests that a given E2F may preferentially regulate only a subset of potential E2F target genes in vivo. This is also consistent with the differential effects of mutation of pRB or p107 and p130 on E2F target gene expression, and with the different biological effects of individual E2F genes. Unraveling which target genes are regulated in which ways (repression, activation, or both) by which E2F genes, and how this specificity is conferred, remains the subject of intensive study.
H. Summary

The E2F transcription factors participate in the regulation of a variety of cellular processes, including progression through G1 and apoptosis. Although there are similarities among the E2F proteins, they are differentially regulated during development and the cell cycle. They also have distinct functions, as determined by both overexpression studies and mutagenesis studies as well as other data. As a family of transcription factors, E2F proteins are believed to effect these distinct functions through the regulation of specific target genes. Indeed, mutagenesis of the E2F inhibitors pRB or p107 and p130 derepresses distinct sets of E2F target genes, and ectopic expression of individual E2F proteins has differential effects on specific E2F target genes. The mechanism permitting this specificity of target gene activation by individual E2F proteins remains unknown. One hypothesis, that E2F proteins might vary in their DNA-binding specificity, is explored in CHAPTER II.

The five E2F proteins described in this chapter are known to be insufficient to account for all of the cellular E2F DNA-binding activity. CHAPTER III describes the identification of a novel member of the E2F family, E2F-6, and its initial characterization. CHAPTER IV details the characterization of a zinc-finger protein interacting specifically with E2F-6. Finally, CHAPTER V discusses the impact of these findings on the larger questions of E2F function and regulation.
I. References


Hu, Q. J., Dyson, N., and Harlow, E. (1990). The regions of the retinoblastoma protein needed for binding to adenovirus E1A or SV40 large T antigen are common sites for mutations. Embo J 9, 1147-55.


CHAPTER II

The individual E2F family members bind a common DNA recognition sequence

B Fairchild, S Estes and JA Lees

(Author’s contribution: all data, except some of the sequencing associated with the binding site selections.)

A. Abstract

The E2F transcription factors have been implicated in the regulation of cellular proliferation and of apoptosis. They are inhibited by the product of the retinoblastoma-susceptibility gene, pRB, and by its homologs p107 and p130. E2F transcription factors are heterodimers composed of an E2F subunit and a DP subunit. In humans, six E2F proteins and two DP proteins have been identified; the pRB family members bind to specific subsets of the E2F proteins. Both overexpression studies and genetic studies indicate that the E2F proteins have a variety of distinct functions in vivo. Mutation of pRB family members derepresses distinct subsets of E2F target genes, as does overexpression of individual E2F proteins. These observations strongly suggest that the E2F proteins regulate different subsets of E2F-responsive genes, and raises the possibility that E2F proteins could differ in their DNA binding specificities. We have determined that the preferred DNA binding sequences for complexes containing DP-1 with E2F-1, -2, -3, or -4 are indistinguishable. Each sequence selected for binding to one of the complexes also binds the others with comparable affinities. More importantly, each sequence in the context of a minimal promoter can serve as a site for transcriptional activation by any of the complexes. Sequence differences within E2F sites are therefore unlikely to mediate target specificity by particular E2F proteins in vivo.
B. Introduction

E2F is a transcriptional regulator that plays a pivotal role in the regulation of cellular proliferation (reviewed in Beijersbergen and Bernards, 1996; Dyson, 1998). Many E2F-responsive genes have been identified and their products are components of either the DNA replication (for example DNA polymerase α, HsOrc1, and cdc6) or the cell cycle control (for example cyclin E, cyclin A and cdc2) machinery. In each case, E2F acts to restrict the transcription of these genes to the stage of the cell cycle in which their products are known to function.

E2F is regulated by the retinoblastoma protein (Bandara and La Thangue, 1991; Chellappan et al., 1991), a tumor suppressor that is functionally inactivated in almost all human tumors (reviewed in Dyson, 1998). Consistent with its anti-proliferative role, the binding of the retinoblastoma protein (pRB) blocks the transcriptional activity of E2F (Hiebert et al., 1992). In addition, overexpression studies indicate that the resultant pRB•E2F complex can act as a transcriptional repressor in which E2F provides the sequence-specific DNA binding and pRB inhibits transcription through the recruitment of a histone deacetylase (reviewed in DePinho, 1998). The interaction between pRB and E2F is regulated by the cell cycle-dependent phosphorylation of the retinoblastoma protein by one or more of the cell cycle-dependent kinases (reviewed in Dyson, 1998). Phosphorylation of the retinoblastoma protein releases it from E2F, switching E2F-responsive genes from a fully repressed to a fully induced state. Consistent with this model, the timing of E2F-responsive gene transcription correlates closely with the timing of pRB phosphorylation at the G1/S transition.

E2F is also regulated by two additional proteins, p107 and p130 (Cao et al., 1992; Devoto et al., 1992; Shirodkar et al., 1992; Cobrinik et al., 1993). These two proteins share significant sequence similarity with the retinoblastoma protein (Ewen et al., 1991; Hannon et al., 1993; Li et al., 1993; Mayol et al., 1993) and they repress E2F in a similar manner to pRB in overexpression assays (Zhu et al., 1993; Starostik et al., 1996). Despite these similarities, the loss of these "pocket proteins" has dramatically different biological effects in mice. Mice heterozygous for a loss-of-function mutation at the retinoblastoma locus develop tumors at 100% penetrance, whereas mice
lacking only p107 or p130 do not demonstrate an increased tumor predisposition (Jacks et al., 1992; Cobrinik et al., 1996; Lee et al., 1996). Homozygous mutations at these three loci also have different developmental effects in mice (Jacks et al., 1992; Cobrinik et al., 1996; Lee et al., 1996; LeCouter et al., 1998a; LeCouter et al., 1998b). Thus, despite the similarities among pRB, p107, and p130, they have distinct biological functions *in vivo*.

The pocket proteins also appear to regulate E2F in distinct ways. First, pRB, p107 and p130 bind to E2F at distinct stages of the cell cycle (Mudryj et al., 1991; Cao et al., 1992; Lees et al., 1992; Shirodkar et al., 1992; Chittenden et al., 1993; Cobrinik et al., 1993; Moberg et al., 1996). Second, the loss of pRB, p107 and/or p130 in otherwise isogenic cell lines results in the specific deregulation of different subsets of E2F-responsive genes (reviewed in Dyson, 1998). These findings have led a number of labs to suggest that the pRB•E2F, p107•E2F and p130•E2F complexes will have distinct target specificities *in vivo*.

To date, at least eight human genes have been identified that encode components of E2F DNA-binding activity (reviewed in Dyson, 1998). These can be divided into two distinct groups, termed E2F (1 through 6) and DP (1 and 2), whose products heterodimerize to generate functional E2F complexes (Bandara et al., 1993; Helin et al., 1993; Krek et al., 1993; Wu et al., 1995). In *vivo* studies indicate that the endogenous E2F activity arises from the concerted action of multiple E2F/DP complexes (Wu et al., 1995; Moberg et al., 1996). Although these complexes do share, to some extent, the ability to promote cell cycle progression when ectopically expressed, both overexpression studies and genetic studies indicate that these complexes also have distinct functions. Ectopic expression of E2F-1, for example, can induce apoptosis in a rat fibroblast cell line (REF52), whereas the overexpression of E2F-2, -3, -4, or -5 does not (DeGregori et al., 1997). Similarly, loss of E2F-1 in the mouse yields a defect in thymocyte apoptosis (Field et al., 1996), whereas loss of E2F-5 affects choroid plexus function (Lindeman et al., 1998). A specific role for E2F-3 has been proposed in the G1/S transition in cycling cells (Leone et al., 1998). E2F complexes therefore have individual functions in addition to any shared ones.
pRB, p107 and p130 each target a subset of the endogenous E2F/DP complexes (Dyson et al., 1993; Lees et al., 1993; Beijersbergen et al., 1994; Ginsberg et al., 1994; Hijmans et al., 1995; Sardet et al., 1995; Vairo et al., 1995). Although the DP subunit is essential for pocket protein binding, the specificity of this interaction is determined by the E2F moiety. E2F-1, -2, or -3 complexes bind specifically to pRB. In contrast, p107 and p130 only associate with E2F-4 and -5 complexes, and E2F-6 does not associate with any of the pocket proteins (Cartwright et al., 1998; Gaubatz et al., 1998; Trimarchi et al., 1998). Sequence comparisons also clearly differentiate the pRB-specific E2F proteins (E2F-1, -2, and -3) from the others. Strikingly, in the DNA-binding domain there are several residues that are absolutely conserved among the pRB-specific E2F proteins, whereas a different amino acid at that position is conserved between the p107/p130-binding E2F proteins. These observations have led several groups to propose that the E2F proteins will regulate different E2F-responsive genes, and that this would account for the different biological properties of the pocket proteins and of the E2F genes. Indeed, ectopic expression of individual E2F proteins has been demonstrated to differentially activate E2F-responsive genes in multiple cell lines (DeGregori et al., 1997; Dirks et al., 1998). For example, expression of E2F-1 efficiently activates the expression of cyclin A but not of cdk2 in REF52 cells, whereas the opposite is true for E2F-3, and E2F-2 is the most efficient at activating DHFR expression (DeGregori et al., 1997). These data argue very strongly that individual E2F genes regulate specific E2F-responsive promoters \textit{in vivo}.

E2F DNA-binding activity was originally defined as the ability to recognize and bind to one of the E2F sites (TTTCGCGC) derived from the adenoviral E2 promoter. However, E2F-responsive elements have been identified within many cellular promoters and there is considerable variation between these sequences and the canonical E2 site. Moreover, a single nucleotide difference between two known E2F-response elements has been proposed to alter the spectrum of associated E2F complexes (Wells et al., 1996; Wells et al., 1997). This suggests that differences in the target specificity of the two E2F subgroups could arise, at least in part, from differences in their DNA binding specificity. In this study, we have used an \textit{in vitro} DNA binding site selection
assay to identify the optimal DNA binding sequences of four different E2F/DP-1 heterodimers. The high affinity DNA binding sites selected for these complexes were indistinguishable from one another. All four E2F/DP-1 complexes bound to these sites with similar, high affinities, and transactivated a reporter bearing these sites with similar efficiencies. We therefore conclude that variations in the primary sequences of high-affinity E2F sites are unlikely to specify which E2F proteins can regulate a given promoter.
C. Results

C1. Production of recombinant proteins

To better understand the DNA binding specificity of the individual E2F/DP complexes, we sought to identify their optimal recognition sequences. To establish a convenient source of each complex, we used baculoviral vectors to express the proteins in insect cells. A baculovirus-based system provided several advantages over other expression systems (Dynlacht et al., 1994). First, it allows the generation of insect cell extracts containing high levels of the desired protein. Second, it confers the ability to form a functional E2F complex by tolerating the infection of both E2F- and DP-expressing viruses. Finally, the E2F complexes expressed in this system are likely to contain any post-translational modifications necessary for their normal function. To take advantage of this system we constructed a series of recombinant baculoviruses expressing hemagglutinin (HA)-tagged E2F-2, -3, or -4 (BV-HAE2F-X) or DP-1 (BV-HADP-1). A recombinant baculovirus expressing untagged E2F-1 (BV-E2F-1) was also used (Dynlacht et al., 1994). Single infections of insect cells with these viruses resulted in high level expression of the expected E2F or DP-1 as judged by western blotting of whole cell extracts from infected cells (Figure 1A). BV-HAE2F-4 infection resulted in the expression of multiple species of E2F-4, consistent with the multiple phosphorylated forms of the protein found in mammalian cells (Figure 1A and data not shown).

We next asked whether the resulting E2F proteins were functional by testing their ability to dimerize with DP-1 and to bind to DNA. Insect cells were dually infected with BV-HADP-1 and individual E2F viruses and the resulting cell extracts were tested for DNA binding activity in gel shift assays using an oligonucleotide containing the canonical E2 site as a probe. At the levels of input protein used, we were unable to detect any E2F DNA-binding activity in extracts of cells that had been mock- or singly-infected (Figure 1B, lanes 1, 6, 11, 16 and 21, and data not shown). As expected, extracts of dually infected cells showed considerable levels of E2F DNA binding activity. The presence of the expected E2F subunit and DP-1 in each infection was confirmed by
Figure 1A.

![Image of gel electrophoresis showing bands labeled αE2F-1, αE2F-2, αE2F-3, αE2F-4, and αHA across different lanes labeled MOCK, E2F-1, E2F-2, E2F-3, E2F-4, and DP-1.]

Figure 1B.

![Image of gel electrophoresis showing bands labeled αE2F-1, αE2F-2, αE2F-3, and αE2F-4 across lanes 1 through 25, with lane 1 labeled no antibody.]
Figure 1. (preceding page) Overexpression of E2F subunits by infection of Hi5 insect cells with BV-E2F-1, BV-HAE2F-2-4 and/ or BV-HADP-1. (A) Three days post-infection cell extracts were prepared from singly- or mock-infected cells and analyzed by Western blotting with E2F- or HA-specific monoclonal antibodies. The amounts of protein loaded were as follows: E2F-1, 20 ng; E2F-2, 200 ng; E2F-3, 5μg; E2F-4, 100 ng; DP-1, 500 ng. In each case, an equal amount of mock-infected extract was run in parallel. (B) Generation of high levels of E2F DNA-binding activity by dual infection of Hi5 cells with BV-HADP-1 and a BV-E2F. Extracts were prepared three days post-infection and used in gel shift assays with the canonical E2 site. 50 ng of total protein from each infection were used per binding reaction, except for the BV-HAE2F-3/ BV-HADP-1 infection in which 400 ng were utilized.
the ability of specific monoclonal antibodies to supershift the relevant complexes (Figure 1B lanes 7, 13, 19, and 25, and data not shown).

C2. E2F site selection

To identify DNA binding sites, the E2F complexes were incubated with a pool of 32P-labeled double-stranded oligonucleotides containing an internal region of 25 random positions. Free and bound oligonucleotides were separated by electrophoresis on 4% native gels (Figure 2). To insure that only relatively high affinity E2F binding sites would be captured, a titration of protein concentration was performed and only the lowest input of protein yielding a detectable band in the gel shift was used for the selection. The oligonucleotides bound each E2F complex were isolated from the gel, amplified by PCR and subjected to five additional rounds of selection with a protein titration performed at each round. After the sixth round, the isolated oligonucleotides were cloned and 28-36 independent clones were sequenced for each E2F/DP-1 complex.

Figure 3 shows an alignment of the oligonucleotides sequenced after six rounds of selection with the E2F-1/DP-1 extract. We found that several selected oligonucleotides contained perfect matches with known E2F-responsive elements, validating the efficiency of the selection protocol. These include E2F-response elements from the E2F-1 (1.6), CDC2 (1.22), and DHFR (1.24, 1.28, and 1.31) promoters. The remaining selected oligonucleotides each contained a sequence that was highly homologous to known E2F binding sites. These selected E2F-1 binding sites had two common features. First, these sequences contained an internal G/C-rich core that was usually five or six nucleotides long. Most notable was the presence of a highly conserved CGC trinucleotide at positions 3-5 of the core. Second, the G/C-rich core was frequently flanked on either side by adenine or thymine residues. A comparison of these sequences yielded the consensus sequence WTTSSCGCSWWWW where W represents A or T and S represents G or C (Figure 4).

The oligonucleotides derived from the other E2F/DP-1 selections showed considerable similarity to those derived from the E2F-1/DP-1 selection (Figures 3 and 4). Each of the
Figure 2.

E2F

DP-1

GGCTGAGCTGCTGAACGGATCCNNNNNNNNNNNNNNNNNNNNNNNNNCCTCGAGACTGAGCGTCG
CCGACTCAGACTTGGCTAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGGAGCTGACTCGAGC

Amplify by PCR
Repeat EMSA

Amplify by PCR
Repeat EMSA
Figure 2. (preceding page) Summary of the site selection scheme. Extracts of infected cells were incubated with a pool of randomized oligonucleotides and separated by gel shift assay. After autoradiography, DNA from the shifted complexes was excised from the gel, eluted, and amplified by PCR. A total of six rounds of selection were performed.

Figure 3. (following page) Oligonucleotides derived from six rounds of E2F-1/DP-1 selection. Sequences are aligned to the highly conserved CGC (boldface). Randomized positions within the oligonucleotides are capitalized, whereas invariant positions are depicted in lower case.

Figure 4. (second following page) Comparison of the consensi derived from the selections with DP-1 and E2F-1, -2, -3, or -4. For each position, the frequency at which a given nucleotide appears is indicated.
Figure 3.
Figure 4.

E2F-1/DP-1 Consensus:

E2F-2/DP-1 Consensus:

E2F-3/DP-1 Consensus:

E2F-4/DP-1 Consensus:
sequences selected for binding to E2F-2, -3, or -4 containing complexes contained the CGC trinucleotide. This was typically present in a five or six nucleotide G/C core flanked by A/T-rich sequences. Occasionally, the G/C run extended to seven or more residues in length. However, this extended G/C-rich sequence was detected at similar frequency in each of the individual E2F selections. For all four E2F proteins, there was a marked preference for thymine in the two positions directly 5’ of the G/C core. The nucleotides following the G/C run are enriched in adenine and thymine, and at a significant frequency this A/T-rich stretch is composed principally or exclusively of either adenine or thymine. We therefore conclude that heterodimers of E2F-1, -2, -3, or -4 with DP-1 selectively bind to the same high affinity sequence, WTTSSCGCSWWWW, in vitro.

The trinucleotide CGC was invariant in the selections involving E2F-2, -3, or -4. In contrast, 3 of the 36 E2F-1-derived sequences contained a CGG at these positions. This raised the possibility that E2F-1 complexes may bind with high affinity to sequences containing a CGG. If so, while promoter sequences with a CGC might be regulated by any of the E2F proteins, sequences with a CGG in its place might be regulated exclusively by E2F-1. To test this possibility, we compared the affinities of E2F-1 and other E2F proteins for CGG-bearing sequences (Figure 5). When used in binding competition experiments, oligonucleotides with a CGG sequence bound less well to each of the four E2F complexes than the canonical, CGC-containing E2F sequence. Thus, the CGG trinucleotide promotes only low affinity binding. Furthermore, the oligonucleotides with the CGG sequence were no more efficient at competing for E2F-1 binding than at competing for binding with the other E2F complexes, arguing that E2F-1 complexes are no more tolerant of this sequence variation than other E2F complexes. We therefore conclude that replacement of the CGC trinucleotide with CGG gives rise to a low affinity E2F binding site that does not show any preference for binding to E2F-1 relative to the other E2F family members.
Figure 5.

Transfected protein

<table>
<thead>
<tr>
<th>E2F-1/DP-1</th>
<th>E2F-2/DP-1</th>
<th>E2F-4/DP-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>no competitor</td>
<td>+ E2-CGGG</td>
<td>+ E2-CGGC</td>
</tr>
</tbody>
</table>

Shifting E2 probe

- 0
- 1000
- 2000
Figure 5. (preceding page) E2F complexes share similar, low tolerances for CGG sequences. C33-A cells were transfected with constructs expressing DP-1 and the indicated E2F protein. 1 μg of transfected cell extract was subjected to gel shift analysis. in the presence or absence of 100 ng of the indicated competitor DNA sequences. The amount of binding to the E2-derived probe in the absence or presence of 100 ng of the indicated competitor DNA sequences was determined by quantitation on a STORM 860 (Molecular Dynamics).
C3. Relative affinities of E2F complexes for selected sites

To confirm that any E2F heterodimer could indeed bind to these selected sites, eight sequences, varying in the length and composition of the internal G/C core as well as the arrangement of flanking sequences, were used in gel shift assays. In these experiments the E2F complexes all bound to any given probe with qualitatively similar efficiencies (data not shown). Furthermore, complexes containing DP-2 (instead of DP-1) also bound all eight sequences with relatively high affinity (data not shown), demonstrating that the DNA sequences recognized by DP-1 can also be recognized by DP-2.

In order to define more precisely the affinity of the different E2F complexes for these selected sites, we selected one each of the E2F-1-, E2F-2- and E2F-4-selected sequences to be used as competitors for the binding of the E2F/DP-1 dimers to the canonical E2F site (Figure 6). Although there was as much as a ten-fold difference in the affinity of the E2F heterodimers for different selected sites (Figure 6, compare 3P with 4E), the relative affinity of the E2F complexes for any given site did not vary more than two- to four-fold. This confirms that, in vitro, subtle sequence differences within or outside an E2F site do not selectively dictate which family members can bind to that site.

C4. Sequence variations do not confer E2F complex specificity in transactivation

Based on the in vitro binding studies, we hypothesized that variations in the sequences of E2F sites are unlikely to determine which E2F proteins are capable of promoting transcription at a given promoter. To test this notion, we constructed a series of reporter constructs that contained a tandem repeat of the three E2F DNA-binding sites, 1.30, 3P and 4E, upstream of a common minimal promoter. Since the transcriptional activity of the individual E2F proteins is known to vary considerably (Verona et al., 1997), we compared the ability of each E2F complex to activate the test reporters versus one containing the canonical E2F site from the E2 promoter. We saw a significant difference in the activity of these test reporters (Figure 7). The reporter containing the low affinity E2F binding site, 3P, gave the lowest levels of activation in the reporter assays.
Figure 6.

![Graph showing relative affinity for different competitors: E2F-1/DP-1, E2F-2/DP-1, E2F-3/DP-1, E2F-4/DP-1. The x-axis represents different competitor concentrations, and the y-axis shows the relative affinity.](image-url)
Figure 6. (preceding page) Competition for recombinant E2F DNA-binding with selected sites. 50 ng (E2F-1 and 2), 200 ng (E2F-3) or 100 ng (E2F-4) of extract from dually infected cells were incubated with a 32P-labelled double-stranded wild-type E2 oligonucleotide in the presence or absence of increasing amounts of the indicated unlabelled competitor (2, 10, 50, or 300 ng).

Figure 7. (following page) E2F-mediated transactivation through selected sites. C33-A cells were transfected with an artificial reporter construct containing two copies of a wild-type or mutant E2F site, or two copies of selected sequences 1.30, 3P, or 4E, upstream of a CAT reporter; CMV-based constructs expressing DP-1 and the indicated E2F were also transfected. For E2F-1, 100 ng of CMV-E2F-1 and 100 ng of CMV-DP1 were transfected; for E2F-2 and -3, 300 ng of the CMV-E2F construct and 300 ng of CMV-DP-1; for E2F-4, 1000 ng of CMV-E2F-4 and 1000 ng CMV-DP-1; the total amount of CMV promoter present in each transfection was maintained at 2 μg by cotransfection of an appropriate amount of CMV plasmid. Depicted in the graph is the ratio of the activity of each E2F/DP complex on the reporter in question to its activity on the wild-type reporter. Transfection efficiency was internally controlled by the cotransfection of 2μg of RSV-luciferase in each experiment. The error bars depict the standard error of the mean.
Figure 7.

[Bar chart showing relative transcriptional activity for different reporters (E2, E2m, 1.30, 3P, 4E) across different conditions (E2F-1/DP-1, E2F-2/DP-1, E2F-3/DP-1, E2F-4/DP-1)].
contrast, the presence of the higher affinity E2F binding sites, 1.30 or 4P, resulted in significantly higher reporter activity. Most importantly, when compared to their ability to activate the canonical E2 reporter, we did not detect any significant difference in the ability of the individual E2F complexes to activate a given test reporter (Figure 7). This was observed over a range of input plasmid concentrations (data not shown), proving that the reporter system is not at saturation. We therefore conclude that the sequence of the E2F site can significantly alter the efficiency of E2F binding and thereby the level of transactivation but it is unable to specify the identity of the E2F complexes which can bind and regulate the target gene.
D. Discussion

The E2F genes vary both in their biological functions and in their ability to activate the expression of endogenous target genes when overexpressed. In this paper, we have investigated the possibility that these differences in E2F function could result from differences in their DNA binding specificities. Through an in vitro site selection assay we have identified high-affinity binding sites for heterodimers containing human DP-1 and E2F-1, -2, -3, or -4. The high-affinity sites for each of these complexes showed considerable homology not only to each other but also to known E2F binding sites from E2F-responsive promoters. By comparing the oligonucleotides selected in these assays, we propose the following consensus E2F DNA-binding site:

WTTSSCGCSWWW, where W represents A or T and S represents G or C. Beyond these thirteen positions, we detected no obvious nucleotide preference at any position, suggesting that flanking sequences do not contribute to binding site specificity in these in vitro assays. Based on DNA binding site competition studies, the internal CGC trinucleotide appears to be critical to high-affinity DNA binding. This motif is not, however, sufficient for binding: it is always found in the context of a longer series of guanine and cytosine residues. High affinity E2F binding seemed to tolerate some flexibility in the length of this G/C run. Although G/C stretches of five, six, or seven residues were most abundant in the selected sites, examples with eight or nine residues were occasionally observed. Regardless of the length of the G/C core, each of the E2F complexes tested showed a marked preference for adenine or thymine at the positions directly flanking the G/C-rich sequences. This was most obvious in the first two positions prior to the core which were strongly biased toward thymine.

The similarity of the sequences derived from the E2F-1, -2, -3, and -4 selections suggested that there may be little or no difference among high-affinity DNA binding sites recognized by the various E2F complexes. Indeed, binding competition studies with sequences derived from the E2F-1, -3, and -4 selections suggest that complexes with DP-1 and E2F-1, -2, -3, or -4 can all bind each of these sequences with comparable affinities. This is consistent with studies comparing
the affinities of various E2F complexes for a sequence from the murine B-myb promoter (TCTCCGCCAAGT) to their affinities for the canonical E2 sequence. Complexes containing E2F-1, -2, -3, -4, or -5 have been reported to bind those two sequences with similar affinities, in the presence of murine DP-1, -2, or -3 and in the absence or presence of pRB or p107 (Zwicker et al., 1996). In both that study and in the present one, the largest differences in the affinities of individual E2F complexes for a given DNA sequence were two- to four-fold. Thus, sequences that bind to one E2F complex with high affinity can also bind to other E2F complexes with similar affinities. Furthermore, the DNA sequences from the binding site selection did not differentiate between E2F-1, -2, -3, or -4 in a transcriptional activation assay. We therefore conclude that the differences in the effects of specific E2F proteins at individual promoters are unlikely to be mediated by differences in the sequences of the E2F DNA binding sites.

Tao et al. have previously reported that the consensus binding sites for E2F-4/DP-1 and for pRB/E2F-1/DP-1 complexes differ from those of E2F-1/DP-1, E2F-1/DP-2, and E2F-4/DP-2 (Tao et al., 1997). This conclusion is clearly at odds with our own. We suggest that this discrepancy is due to the manner in which they derived their consensus binding sites. First, their analysis included multiple copies of a relatively small number of identical oligonucleotides, suggesting that they resulted from repeated PCR amplification of a small number of selection events. Although each of our selections yielded twenty-eight to thirty-six inserts per experiment, every sequence was different suggesting it arose from an independent binding event. Second, Tao et al. included partially palindromic sites in both possible orientations. In each case, we believe that the inclusion of these repeat sequences is likely to have introduced significant bias into the consensus. In both the E2F-4/DP-1 and the pRB/E2F-1/DP-1 selections there were high percentages of identical sequences, most of which were palindromic. Indeed, if the identical sequences from their E2F-4/DP-1 selection are excluded, and the palindromic sequences are only counted once, the resulting consensus strongly resembles their E2F-1/DP-1 consensus, and our consensus for both complexes. This reaffirms our observations that different E2F proteins share the same preferred DNA recognition sequences.
There are also other important similarities between our data and theirs. For example, we both find a tolerance for either cytosine or guanine at a variety of positions in the C/G core (Tao et al., 1997). Also, just as we find that affinity differences among E2F complexes for our selected sequences do not vary by more than two- to four-fold, they found that the ratio of the half-life of E2F-1/DP-1 to the half-life of E2F-4/DP-1 complexes on individual sequences does not vary by more than two- to four-fold. Indeed, they found a four-fold difference between two sequences that do not vary in their E2F binding site (TTTTCGCGCCAAGGA and TTTTCGCGCCAAAAGGG), suggesting that this may represent the degree of variation within the assays. They also found that their consensus sequences varied in how efficiently they confer transcriptional repression when substituted for the E2F site in the B-myb promoter. Indeed, most of the differences they observe in these reporter assays could be explained by differences in the efficiency of the sites in mediating the transcriptional effects of E2F, as we observed in our assays. Other factors may also contribute to the differences they observe: the introduction of consensus sequences that are extended palindromes might affect the promoter architecture, or unrelated complexes could fortuitously recognize their sequences.

E2F complexes have different biological roles in the organism. E2F-5, for example, has a specific, required function in the choroid plexus (Lindeman et al., 1998). E2F-1 promotes apoptosis (reviewed in Dyson, 1998), whereas E2F-3 may have a specific function in S phase entry in cycling cells (Leone et al., 1998). E2F complexes also seem to differentially regulate E2F-responsive genes (DeGregori et al., 1997; Dirks et al., 1998). This study has focused on sequences selected for their high-affinity binding to E2F complexes, and therefore has not addressed the possibility that certain low-affinity E2F sites may preferentially bind particular E2F complexes. Variations in the sequences of high-affinity E2F DNA binding sites, however, do not appear to confer specificity of E2F complex binding or of transactivation.

We propose that this specificity may instead result from differences in the promoter context of the E2F sites. This is consistent with observations that, although the B-myb E2F site specifically mediates transcriptional repression in G0/G1 and the DHFR site specifically mediates
transcriptional activation at the G1/ S transition, these effects do not depend on the sequence of the E2F site, but on the promoter (Fry et al., 1997). Some promoters may contain transcription factors that can selectively bind and stabilize particular E2F proteins, conferring target specificity. The transcriptional activation domains of individual E2F proteins may differ in their ability to promote transcription, depending on the identity of other factors at the promoter. Interactions of other transcription factors with DP or pRB family members may also prove important. Indeed, all of these factors may participate in the specificity of E2F transcriptional regulation in vivo.
E. Materials and Methods

Generation of recombinant E2F

The baculovirus transfer vector pVL1392 (Pharmingen) was modified by adding the coding sequence of the hemagglutinin (HA) tag as an EcoRI/BamHI restriction fragment to create the plasmid pVL1392-HA. The coding sequences of E2F-2, -3, -4 and DP-1 were subsequently cloned as BamHI restriction fragments into this vector. Recombinant baculoviruses were generated in Sf9 cells grown in Grace’s media supplemented with 10% foetal bovine serum according to the manufacturer’s instructions. A baculovirus expressing E2F-1 was generously provided by Brian Dynlacht (Dynlacht et al., 1994). After three rounds of amplification, titers were determined and Hi5 cells were singly or dually (E2F/DP-1) with an MOI of 5. Three days post-infection, cells were harvested and whole cell extracts prepared as previously described (Moberg et al., 1996), except that the lysates were cleared by centrifugation at 14,000 RPM for one hour.

Western blotting

Extracts from singly infected Hi5 cells were separated by SDS-PAGE (10%) and electroblotted to a nylon membrane. Total protein loaded varied from 20 ng (BV-E2F-1 infection) to 5 μg (BV-HAE2F-3 infection) with an equal amount of uninfected Hi5 extract run in parallel as a control for each infection. The following monoclonal antibodies were used for detection: KH20 (E2F-1) (Helin et al., 1992), LLF2-4/ LLF2-7 (E2F-2), LLF3-1 (E2F-3) (Moberg et al., 1996), LLF4-2 (E2F-4) and 12CA5 (DP-1). Antibodies LLF2-4, LLF2-7, and LLF4-2 were generated as previously described (Moberg et al., 1996). Antigen/antibody complexes were detected by enhanced chemiluminescence (Amersham).
Site selections

A pool of radioactive, double-stranded oligonucleotides 63 residues in length with 25 randomized positions flanked by constant regions was incubated with insect cell extracts containing heterodimers of DP-1 with E2F-1, -2, -3, or -4. DNA that was bound to proteins was separated from unbound DNA by gel shift as described below. This protein-bound DNA was eluted from the gel, amplified by PCR in the presence of $\alpha$-$^{32}$P-dCTP, and gel purified as described elsewhere (Pomerantz and Sharp, 1994). This PCR product was then incubated with E2F/DP-1 complexes and the entire procedure repeated another five times. At all stages, only the minimum amount of insect cell extract necessary to produce a gel shift was included. For E2F-1/DP-1 complexes, the amounts of cell lysate used in the selections were 50 ng of total protein, 50 ng, 50 ng, 2.5 ng, 0.5 ng, and 0.5 ng for rounds 1 through 6, respectively; for E2F-2/DP-1, 50 ng, 50 ng, 50 ng, 2.5 ng, 2.5 ng, and 0.5 ng; for E2F-3/DP-1, 200 ng, 200 ng, 200 ng, 10 ng, 10 ng, and 2.5 ng; for E2F-4/DP-1, 200 ng, 50 ng, 10 ng, 10 ng, 10 ng, and 10 ng. Selected sites were amplified by PCR, cloned into pBluescriptIISK+ (Stratagene) and characterized by a double-stranded dideoxynucleotide sequencing protocol (USB). Sequences were aligned with assistance from the CLUSTAL program found at http://dot.imgen bcm.tmc.edu:9331/multi-align/Options/clustalw.html.

Gel shift assays and affinity studies

Gel shifts were performed as described (Moberg et al., 1996) with extracts from cells infected with baculoviruses expressing DP and E2F subunits. For supershifting E2F/DNA complexes the following monoclonal antibodies specific for the indicated E2F were included in the binding reaction: KH20 (E2F-1) (Helin et al., 1992), LLF2-5 (E2F-2), LLF3-1 (E2F-3) (Moberg et al., 1996), and LLF4-1 (E2F-4) (Verona et al., 1997). Antibody LLF2-5 was generated exactly as previously described (Moberg et al., 1996). The sequences of the double-stranded oligonucleotides used in the binding site competition experiments are as follows: wild-type E2: ATTTAAGTTTCGCGCCCTTTCCAA; mutant E2: ATTTAAGTTTCGATCCCTTTCCAA; E2-
CGGC: TGACTAGTTTCCCGCTTTCTC; E2-CGGG: TGACTAGTTTCCCGGTTTCTCAAA; 1.30: TCGAGGGTGCGCTTTCGGCGCAATGGATC; 3P: TCGAGGCGC-CAACCCGTTTCCGCGTGTCACGGATC; 4E: TCGAGGTGAAAAGTATTCCCGCGCTTTG- GGATC. Binding quantitation was performed on a STORM 860 using ImageQuant v1.2 (Molecular Dynamics).

Transient transfections and transcriptional reporter assays

C33-A cells were maintained in DMEM supplemented with 10% fetal bovine serum; transient transfections were performed as described (Trimarchi et al., 1998). For reporter assays, chloramphenicol acetyltransferase (CAT) reporter constructs were generated by replacing the XhoI/SalI region of E1b-CAT with two tandem copies of the following sequences: for 2xE2-CAT, GAAGTTTCGCGCCCTTTG; for 2xE2mut-CAT, GAAGTTTCGATCCCTTTG; for 2x1.30-CAT, GAAGTTTCGCGCCCTTTG; for 2x3P-CAT, GAAGTTTCGCGTGTCG; for 2x4E-CAT, GAAGTTTCGCGTGTCG. 4 μg of one of these reporters was cotransfected with 2 μg of RSV-luc, 19 μg of pBluescriptIIKS+ and the indicated CMV-based E2F and DP-1 expression constructs; the total amount of CMV promoter in each transfection for reporter assays was maintained at 2 μg by the addition of CMV plasmid as appropriate. CAT and luciferase activities were assayed as described (Lees et al., 1993).
F. Acknowledgements

We gratefully acknowledge Brian Dynlacht for providing the E2F-1-expressing baculovirus, many members of the Lees lab, past and present, for helpful advice and discussions, Steve Bell and Patrick McCaw for advice about quantitation, Joel Pomerantz and Ernest Fraenkel for critical help in setting up the site selection experiments, and Raluca Verona and Jeff Trimarchi for critical reading of sections of this manuscript.
G. References


Dirks, P. B., Rutka, J. T., Hubbard, S. L., Mondal, S., and Hamel, P. A. (1998). The E2F-family proteins induce distinct cell cycle regulatory factors in p16-arrested, U343 astrocytoma cells [In Process Citation]. Oncogene 17, 867-76.


accelerated cell cycle in mice lacking the Rb-related p107 gene [In Process Citation]. Mol Cell Biol 18, 7455-65.


CHAPTER III

E2F-6, a member of the E2F family that can behave as a transcriptional repressor

JM Trimarchi, B Fairchild, R Verona, K Moberg, N Andon, and JA Lees

Copyright (1998) National Academy of Sciences, U. S. A.

(Author’s contribution: identification of E2F-6)

A. Abstract

The E2F family of proteins is required to establish the correct cell-cycle-dependent transcription of genes that direct the process of cell division. All previously identified E2F proteins can act in a similar manner; depending on whether or not they are associated with the cell cycle inhibitors the retinoblastoma protein (pRB), p107, or p130, they can either repress or activate the transcription of E2F responsive genes. We now report the cloning and characterization of another E2F family member, E2F-6, whose structure is reminiscent of the dominant inhibitors of other transcription factor families. The dimerization and DNA-binding properties of E2F-6 are similar to those of the other E2F family members. However, it is not regulated by pRB, p107, or p130, and it is unable to activate transcription. Instead, it can act to repress the transcription of E2F responsive genes by countering the activity of the other E2F complexes via a pRB-, p107-, or p130-independent mechanism.
B. Introduction

The retinoblastoma gene (RB-1) is one of the best studied of the tumor suppressor genes. Although it was originally identified by virtue of its absence in retinoblastomas, subsequent studies have shown that it is absent or mutated in at least one-third of all human tumors (Weinberg, 1992). The product of this gene [the retinoblastoma protein (pRB)] is also an essential target of the transforming proteins of the small DNA tumor viruses (Dyson and Harlow, 1992). In untransformed cells, the growth-suppressive properties of pRB are regulated by its cell-cycle-dependent phosphorylation (Bartek et al., 1996). This phosphorylation is catalyzed by the cell-cycle-dependent kinase cyclin D•CDK4/6 (Bartek et al., 1996). Consistent with this hypothesis, many human tumors contain activating mutations within either the cyclin D1 or CDK4 genes or have lost the CDK4/6-specific inhibitor p16 (Pollock et al., 1996).

The growth-suppressive properties of pRB are largely dependent upon its ability to regulate a cellular transcription factor, E2F (Nevins, 1992; Bartek et al., 1996). Many E2F-responsive genes have been identified, and their products are required for entry into, or passage through, the cell cycle. Consistent with its antiproliferative role, pRB inhibits the transcriptional activity of E2F (Hiebert et al., 1992). Moreover, the resultant pRB-E2F complex actively represses the transcription of E2F-responsive genes by blocking the activity of adjacent transcription factors (Weintraub et al., 1992; Bremner et al., 1995; Weintraub et al., 1995). Phosphorylation of pRB causes it to dissociate from E2F, thereby switching E2F-responsive genes from the repressed to the induced state.

E2F is regulated by two additional proteins called p107 and p130 (Beijersbergen and Bernards, 1996). These proteins share significant sequence similarity with pRB (Dyson and Harlow, 1992; Beijersbergen and Bernards, 1996), and in overexpression assays they repress E2F in a similar manner to pRB (Zhu et al., 1993; Starostik et al., 1996). However, genetic studies indicate that pRB, p107, and p130 have distinct properties in vivo. Whereas pRB is mutated in 30% of all human tumors, neither p107 nor p130 has been proven to suppress tumorigenesis...
Moreover, analysis of mutant mouse strains indicates pRB is essential for development, whereas loss of p107 or p130 in a similar genetic background does not alter viability or tumor incidence (Clarke et al., 1992; Jacks et al., 1992; Lee et al., 1992; Cobrinik et al., 1996; Lee et al., 1996). It is unclear exactly how the different biological properties of these proteins relate to the manner in which they regulate E2F, but pRB, p107, and p130 are known to bind E2F at defined but distinct stages of the cell cycle (Beijersbergen and Bernards, 1996).

To date, at least seven human genes have been identified that encode components of the E2F transcriptional activity (Beijersbergen and Bernards, 1996). These can be divided into two distinct groups, named E2F (E2F-1 through -5) and DP (DP-1 and -2). The products of these groups heterodimerize to give rise to high-affinity DNA-binding activity and transcriptional activation (Bandara et al., 1993; Helin et al., 1993; Krek et al., 1993; Wu et al., 1995). In vivo studies confirm that the endogenous E2F activity is generated from the combined properties of multiple E2F-DP complexes (Wu et al., 1995; Moberg et al., 1996). The individual E2F-DP species have different pRB-, p107-, and p130-binding properties. Although the DP subunit is essential, the E2F moiety mediates the specificity of this interaction. Complexes containing E2F-1, -2, or -3 associate with pRB but not p107 or p130 in vivo (Dyson et al., 1993; Lees et al., 1993). In contrast, E2F-4 and -5 complexes are capable of binding p107 and p130 (Beijersbergen et al., 1994; Ginsberg et al., 1994; Hijmans et al., 1995). Consistent with these findings, sequence comparisons suggest that the family of E2F proteins can be subdivided into two distinct subgroups. The pRB-specific E2Fs (E2F-1 through -3) have an extended N-terminal domain that is absent in both E2F-4 and -5. There is also considerable variation in the sequence of DNA-binding, dimerization, and transactivation domains between members of the two E2F subgroups (E2F-1 through -3 versus E2F-4 and -5). These observations have led to the hypothesis that these two subgroups will play distinct roles in vivo that will at least partially account for the different biological consequences of loss of pRB, p107, or p130.

To examine the biochemical and functional properties of the endogenous E2Fs, we had previously developed specific antisera for each of the components of the E2F family (Moberg et
al., 1996). With these reagents, we have been able to demonstrate that the known E2F proteins are unable to account for all of the endogenous E2F-DP DNA-binding activity (Moberg et al., 1996). Specifically, gel shift analysis revealed that an activity exists in gel shift analysis with the same DNA binding specificity as other E2F complexes, but that is not recognized by antibodies to the known E2F proteins. The activity is, however, recognized by anti-DP antibodies. Thus, there exists at least one novel protein that, like known E2F proteins, binds to DP and to an E2F DNA binding sequence. We have therefore sought to identify additional members of the E2F family. In this study, we describe the cloning and characterization of an additional E2F family member, E2F-6. The DNA-binding and dimerization domains of E2F-6 are highly related to the corresponding domains of the previously identified family members, but this protein lacks the sequences necessary for either transactivation or pRB-, p107-, or p130-binding. We conclude that the E2F family contains a third subgroup of proteins whose structure is highly reminiscent of the dominant inhibitors of other transcription factor families.
C. Results

C1. Isolation of cDNAs encoding an E2F family member

At the start of this study, five genes had been identified that encode members of the E2F family of proteins. We have shown previously that these proteins account for a significant proportion of the endogenous E2F-DP complexes, but there must be at least one additional E2F (Moberg et al., 1996). The greatest homology between the known E2F family members maps to the DNA-binding domain. In an attempt to identify additional E2F family members, we searched the EST database for cDNA clones with homology to E2F DNA-binding domains. As expected, we were able to identify ESTs derived from E2F-1, -2, -3, -4, or -5. In addition, this search identified one mouse (GenBank accession no. AA041604/AA050073) and one human (GenBank accession no. AA127210) EST that did not correspond to the known E2Fs. These clones were highly related to one another at the nucleotide level, suggesting that they were mouse and human homologs of the same gene. The core of previously identified E2F DNA-binding domains includes an invariant stretch of fifteen amino acids (QKRRIYDITNVLEGI); this motif was highly, but not completely, conserved in one of the predicted translation products and almost all of the variation between these two clones was in the third base position of this ORF. These data strongly suggested that this was the correct ORF and that these cDNAs were excellent candidates to encode another E2F.

With both the EST database and standard library screening techniques, we identified multiple overlapping cDNAs that encompassed 2,027 bp of the human gene. This sequence starts within the ORF and extends to the poly(A)+ tail. None of the human cDNAs diverged from this assembled sequence. We also identified two noncontiguous mouse cDNAs. The first is a 638-bp clone that shares significant homology with the 3' untranslated region of the human gene. The second includes 216 bp of 5' untranslated region followed by the initiating ATG (as judged by the presence of a good Kozak consensus and an upstream in-frame termination codon) and 536 bp of ORF. Sequence comparison suggested that the human cDNA lacks only a short region (18 bp) of 5' coding sequence (data not shown). The overlapping regions of the human and mouse sequence
are highly conserved at both the nucleotide (83%) and predicted amino acid (92%) level (Fig. 1A). We therefore concluded that we had identified the mouse and human orthologs of another gene, hereafter designated E2F-6, that share significant homology with the known E2Fs.

The previously identified E2F proteins have been divided into two distinct subgroups (E2F-1 through -3 versus E2F-4 and -5) on the basis of differences in their amino acid sequence. The central portion of the E2F-6 protein shared considerable homology with the domains of E2F-1 through -5 that are known to mediate their dimerization (both the leucine zipper domain and region of homology known as the marked box) and DNA-binding properties. Within these domains, many of the residues that are absolutely conserved between E2F-1 through -5 are also maintained in E2F-6. However, at a significant proportion of these conserved residues, E2F-6 had alternative codon usage from that found in the other E2F genes (data not shown). Moreover, when we examined the amino acid positions that are known to distinguish E2F-1 through -3 from E2F-4 and -5, the corresponding residue in E2F-6 rarely fit into either subgroup (see Fig. 1B). The distinction between E2F-6 and the other E2F family members was further underscored by the degree of sequence variation outside of the DNA-binding and dimerization domains. The N-terminal domain of E2F-6 was of intermediate length, relative to those of the two E2F subgroups. More importantly, the E2F-6 protein terminated just 42 amino acids beyond the marked box motif (the C-terminal portion of the dimerization domain). As a result this protein lacks the sequences that are known to mediate either the transcriptional activation of pRB-, p107-, or p130-binding properties of the other E2F proteins. Thus, these findings suggest that E2F-6 represents a third subclass of the E2F protein family that is likely to display distinct properties from the previously identified members.

C2. E2F-6 is widely expressed in vivo.

E2F-1 through -5 are all expressed in a wide variety of tissues (Helin et al., 1992; Kaelin et al., 1992; Ivey-Hoyle et al., 1993; Lees et al., 1993; Beijersbergen et al., 1994; Ginsberg et al., 1994; Hijmans et al., 1995). We were therefore interested to establish the expression pattern of
Figure 1.

A

hE2F6  HEKLPSLLLLDPTEETVRRRCRDPINVEGLLPSKIRINLEDNVQYVSRRMKALKVRRPFDVSLYLYTRKFDLVRSAAPGILDNLKVAT
mE2F6  MSQRTARRQPSLVLPTAQTVRRCRDPINVEGLLPSKIRINLEDNVQYVSRRMKALKVRRPFDVSLYLYTRKFDLVRSAAPACILNLKVAT
hE2F6  KLGVRKRRVYDDNLGIDLVEKSSKNHRGVSDDNSGAVPQQKHKLGEESLDSLAMEDALDELIKDAQQLFLATD зрениеNLKVAT
mE2F6  KLGVRKRRVYDDNLGIDLVEKSSKNHRGVSDDNSGAVPQQKHKLGEESLDSLAMEDALDELIKDAQQLFLATD зрениеNLKVAT
hE2F6  HSIQAFHFGQIVIAVKAPAETRLDVPAPREDVTVHISTENSEPVLCVEVQGTSNKRSREVGTSSSESTHPGPEEEENPQSEELE EVSN*

B

DNA binding domain

Marked box

Leucine zipper

Pocket protein binding
**Figure 1.** (preceding page) E2F-6 is another member of the E2F family. (A) Amino acid sequence comparison of human versus murine E2F-6. (B) Amino acid sequence comparison of the human E2F proteins. Domains responsible for DNA binding, dimerization (leucine zipper, marked box), and pocket protein binding are indicated. Conserved residues are denoted in boldface type.
E2F-6. Initially, we isolated poly(A)+ RNA from the indicated human cell lines and screened them for the presence of the E2F-6 mRNA by Northern blotting using a probe derived from the coding sequence (Fig. 2B). In these and every other cell line examined, this probe hybridized with similar stringency to two distinct messages. To date, each of the E2F-6 cDNAs that we have identified corresponds to a single common transcript. It was therefore unclear why the E2F-6 probe detected two different mRNAs. To address this issue, we rescreened these Northern blots with a probe derived from the 3' untranslated region. This second probe also hybridized to the same two transcripts (data not shown), indicating that they must both contain at least some of the sequences from both the coding and 3' untranslated region of our E2F-6 cDNA. We also examined the expression of E2F-6 in human tissues. For these experiments, we screened a human tissue blot (CLONTECH) with the probe corresponding to the E2F-6 coding region (Fig. 2A). As with cultured cells, we detected two E2F-6 transcripts in every tissue examined. We therefore conclude that E2F-6 will be expressed in most, if not all, cell and tissue types. At this time, we are unable to explain the structural difference(s) between the two mRNAs.

Given the existence of the two E2F-6 transcripts, we could not rule out the possibility that alternative splicing could give rise to protein products with two distinct C-termini, one that corresponds to that encoded by our identified cDNA clones and one that more closely resembles the domain structure of the previously identified E2F proteins. To address this issue, we raised a polyclonal antiserum against the predicted human E2F-6 protein. Control experiments confirmed that the antiserum recognized epitopes throughout the E2F-6 protein but did not cross-react to any of the other known E2Fs (data not shown). When tested on Western blots, the antiserum detected a single 35-kDa protein in all tested cell lines that migrated slightly faster than an HA-tagged version of the E2F-6 protein produced by transient transfection (Fig. 2B and data not shown). We therefore conclude that the E2F-6 protein is expressed in vivo and it exists predominantly in the form predicted by the cDNA clones.
Figure 2.
Figure 2. (preceding page) Expression patterns of E2F-6. (A) Northern blot analysis of poly(A)^+ RNA from the indicated cell lines and tissues (human tissue blot from CLONTECH), screened with a probe derived from the full-length E2F-6 coding sequence. (B) Western blot analysis of extracts derived from C33-A cells transiently transfected with pCMV-HA-E2F-6 (200 ng) and from untransfected U2OS, C33-A, ML-1, and 293 cells (50 μg). The positions of HA-E2F-6 and endogenous E2F-6 proteins are indicated by arrows. The band running beneath HA-E2F-6 is a degradation product of this transfected protein (data not shown). The asterisk denotes a nonspecific band.
C3. **E2F-6 displays low-affinity E2F DNA-binding activity.**

Given the unusual structure of E2F-6, we wished to establish whether this protein retained any of the properties of the known E2F family members. All previously identified E2Fs can bind to either of the human DPs and this heterodimerization is known to be a prerequisite for the high-affinity DNA binding. We therefore initiated our analysis by comparing the ability of E2F-1, E2F-4, or an HA-tagged version of E2F-6 in the presence or absence of the DP (either DP-1 or DP-2) or pocket (pRB, p107, or p130) proteins. The transfectants were labelled with [³⁵S]-methionine and then subjected to immunoprecipitation with the indicated antibodies (Fig. 3A). Although there was considerable variation in the efficiency of the individual transfections, these experiments allowed us to assess the DP- and pocket protein-binding properties of the individual E2F proteins. A monoclonal antibody specific for the HA tag was able to recover either DP-1 or DP-2 in approximately stoichiometric amounts with HA-E2F-6 (Fig. 3A, lanes 18 and 21). Moreover, HA-E2F-6 seemed to bind to these proteins as well as either E2F-1 or E2F-4 (Fig. 3A, compare lanes 18 and 21 with lanes 3 and 11 and lanes 6 and 14). This confirmed that E2F-6 is able to heterodimerize with the DP proteins despite the sequence variation within its presumed dimerization domain. Consistent with previous studies, the E2F-1-DP-1 (Fig. 3A, lanes 4 and 5) and E2F-4-DP-1 (Fig. 3A, lanes 12 and 13) complexes were able to associate with pRB and p107 when overexpressed. In contrast, we did not detect any interaction between the HA-E2F-6-DP complexes any of the three pocket proteins (Fig. 3A, lanes 19 and 20 and data not shown). This finding is consistent with the absence of a pocket protein-binding motif within the E2F-6 protein sequence.

After establishing that E2F-6 can associate with the DP proteins, we next examined the DNA-binding activity of these heterodimers. Transient transfection was used to generate the relevant E2F complexes, exactly as described above, and the resultant cell extracts were measured for total protein content. These were then tested in gel retardation assays with the consensus E2F site and equivalent levels of total protein (Fig. 3B). Consistent with previous studies, E2F-1 and E2F-4 were unable to bind DNA in the absence of a co-transfected DP protein (Fig. 3B, lanes 3
Figure 3.

A

B

anti-EZF-1  anti-EZF-4  anti-EZF-6
Figure 3. (preceeding page) Dimerization and DNA binding properties of E2F-6. (A) C33-A cells were transiently transfected with expression vectors encoding E2F-1, E2F-4, or HA-E2F-6 in the presence or absence of DP (DP-1 or DP-2) or pocket proteins (pRB, p107) and immunoprecipitated with the indicated antibodies. (B) C33-A cells were transiently transfected with the identical combinations of expression vectors as in A. Gel shift assays were carried out on the indicated unlabelled cell extracts (1.5 µg of total protein per lane) by using the consensus E2F site from the adenoviral E2 promoter (TTTCGCGCCCTTT). Western blot assays were carried out on the same unlabelled cell extracts (300 ng of total protein per lane) by using monoclonal antibodies against E2F-1 (KH95), E2F-4 (LLF4-1), or the HA tag (12CA5). Gel retardation assays have also been conducted by using the additional E2F sites TTTCCGCGCTTT, TTTCCGCCCATAA, TTTCCGCGGTGT, or ATTCCCGCGCTTT with similar differences in affinity.
and 8). In a similar manner, we were unable to detect any increase in the levels of E2F DNA binding activity in cells that were transfected with HA-E2F-6 alone (Fig. 3B, lane 13). In the same assay, E2F-1, -4, and -6 were able to bind to DNA when associated with either DP-1 (Fig. 3B, lanes 4, 9, and 14) or DP-2 (Fig. 3B, lanes 7, 12, and 17). However, we consistently detected less E2F DNA binding activity in extracts derived from cells transfected with HA-E2F-6 rather than E2F-1 or E2F-4 (Fig. 3B, compare lanes 14-17 with lanes 4-7 and 9-12). Given this finding, we also assessed the expression levels of these E2Fs in Western blots (Fig. 3B). Because the blots were probed with different antibodies, we cannot make definitive conclusions about the relative levels of these proteins. However, these antibodies have similar avidities for their respective antigens (unpublished observations), suggesting that E2F-1 and E2F-4 are not expressed at significantly higher levels than HA-E2F-6. This suggests that HA-E2F-6 containing complexes have a lower affinity for the consensus E2F site relative to complexes containing either E2F-1 or -4. This low DNA-binding activity was observed with multiple independent transfections and was not improved by removal of the HA tag (data not shown). Given these observations, we also tested the E2Fs for their ability to bind to other E2F binding sites (data not shown). The DNA-binding activity of E2F-6-DP did increase when we used the probes TTTCCCGCC(A/T)(A/T)(A/T). However, this site was previously identified by site selection assays as the preferred recognition sequence of complexes containing E2F-1, -2, -3, or -4, and these species also have a higher affinity for this sequence than for any other E2F site (see CHAPTER II). In fact, relative to the other E2F species, E2F-6-DP complexes bound less well to each of the five probes tested. We therefore conclude that E2F-6 has a lower affinity for DNA, at least when associated with either of the known DP proteins.

**C4. E2F-6 is unable to activate the transcription of E2F-responsive genes.**

E2F-6 lacks the C-terminal sequences that are known to mediate the transcriptional activity of E2F-1 through -5. However, its shortened C-terminal domain is highly charged and contains a significant proportion of acidic residues (22%), suggesting that it might activate transcription via
this alternative motif. To test this hypothesis, we transiently transfected C33-A cells with a chimeric reporter construct, E2F4-CAT, in which the expression of the CAT gene is controlled by a minimal promoter containing four consensus E2F sites upstream of the E1B TATA box (Helin et al., 1992) and increasing amounts of the eukaryotic expression vectors encoding E2F-1, E2F-4, or HA-E2F-6 (Fig. 4A). These transfections were conducted in either the absence or presence of CMV-DP-2 as indicated. E2F-1 and E2F-4 both substantially increased the activation of this reporter. In contrast, E2F-6 did not bring about any increase in the level of E2F transcriptional activity, either in the absence or presence of cotransfected DP proteins (Fig. 4A). Instead, the increasing input levels of E2F-6 steadily inhibited the ability of the endogenous E2F complexes to activate this reporter (Fig. 4B). We did not see any inhibition of reporter activity when the E2F sites were deleted, indicating that this repression was specific for E2F-responsive reporters (data not shown). This inhibition was observed in either the absence or presence of an excess of exogenously expressed DP protein (Fig. 4B and data not shown).

After finding that E2F-6 can block the activity of the endogenous E2F complexes, we wanted to establish whether it could also inhibit a cotransfected E2F. For this experiment, C33-A cells were transiently transfected with the E2F4-CAT reporter in either the absence or presence of CMV-E2F-1 (50 ng) and increasing amounts of CMV-E2F-6 (Fig. 4C). At this input level, E2F-1 increased the activation of this reporter by approximately six-fold. This was effectively inhibited by increasing input levels of CMV-E2F-6. Exactly as described above, this repression was observed in either the absence or presence of cotransfected DP (data not shown). We therefore conclude that E2F-6 is unable to activate transcription and, at least when overexpressed, can inhibit the transcriptional activity of the other E2F species.
Figure 4.

A

- CMV-DP

- CMV-EP

B

- CMV-DP

C

- CMV-DP

Reporters
**Figure 4.** (preceding page) Effects of E2F-6 on transactivation of an E2F₄-CAT reporter C33-A cells were transiently transfected in duplicate with 4 μg of E2F₄-CAT, 2 μg of Rous sarcoma virus-luciferase (as an internal control), 14 μg of carrier DNA (pBIKS⁺ and pCMV-neo-Bam) in the absence or presence of 3 μg of pCMV-HA-DP2 plus pCMV-E2F expression vectors as indicated. CAT and luciferase activity were determined 24 h after transfection. The values shown are the average of duplicate transfectants for representative experiments.
D. Discussion

E2F plays a pivotal role in the regulation of cellular proliferation by controlling the expression of genes that are essential for either entry into, or passage through, the cell cycle. This activity is extremely complex, arising from the combined action of multiple E2F-DP heterodimers. Although there are differences in their cell cycle regulation, the previously identified E2F-DP species can all display the same transcriptional properties. In the presence of an associated pocket protein, they each behave as transcriptional repressors. Once released from the pocket proteins, these same E2F-DP species can activate transcription. Although the DP moiety is required for activity, the transcriptional activation and pocket protein-binding properties of the E2F-DP complexes are entirely dependent upon sequences within the E2F protein.

Herein we describe the cloning and characterization of the human and murine orthologs of a gene that contains the most conserved domains of the known E2F family members, including the core sequences required for DNA binding and dimerization. Functional studies confirm that the human protein can dimerize with either DP-1 or DP-2 and the resultant complexes can bind to DNA in a sequence-specific manner. These findings confirm that this protein is a genuine member of the E2F family, and we have, therefore, named it E2F-6. Despite this designation, our data indicate that this protein functions in a different manner than the other E2F family members. (i) E2F-6-containing complexes bind to DNA significantly less well than the other E2F-DP species. Although we have yet to calculate dissociation constants for these interactions, our preliminary studies suggest that E2F-6-containing complexes recognize the same target sequences as the other E2F species but they bind to these sites with at least a five-fold lower affinity. (ii) E2F-6 lacks the domains that are known to mediate the transactivation and pocket protein-binding properties of the other E2Fs and it is unable to perform either of these functions. (iii) Overexpressed E2F-6 is able to block the transcriptional activity of either cotransfected E2F-1 or the endogenous E2F complexes. Although the mechanism of this repression is unclear, our data indicate that transcriptional inhibition is not due solely to the sequestration of DP from the other E2F family members by abnormally high levels of E2F-6.
Many other transcription factor families include one or more members that behave as transcriptional inhibitors but there is considerable variation in the mechanism by which repression is mediated. The Id proteins of the myogenic basic-helix-loop-helix family, the I-POU protein of the POU domain family, and the CHOP protein of the c/EBP-like family lack the sequences necessary for high-affinity DNA-binding and, therefore, function in a DNA-independent manner (Benezra et al., 1990; Treacy et al., 1991; Ron and Habener, 1992). Instead, they are able to inhibit transcription by binding to the essential heterodimeric partners of the positively-acting family members and forming nonfunctional complexes. There are also several examples of inhibitory proteins that repress transcription in a DNA-binding-dependent manner. In this case, inhibition can be mediated by either exclusion of other family members from the DNA or through the recruitment of cellular factors that actively inhibit transcription. One example of the latter case is the Mad-Max complex that has recently been shown to mediate repression in a sequence-specific manner through its ability to recruit histone deacetylase activity to its target genes (Pazin and Kadonaga, 1997).

Because the E2F proteins function as part of a heterodimeric complex, it is possible to envisage that the repressive effects of E2F-6 could be mediated by any of the repressive effects of E2F-6 could be mediated by any of the mechanisms described above. Although we cannot rule out any of these models, two observations suggest that E2F-6 is unlikely to act solely through the sequestration of the DP proteins. (i) Unlike the other transcriptional repressors that act by this mechanism, E2F-6 retains the ability to bind to DNA. (ii) At least as observed in our overexpression assays, addition of excess DP protein does not prevent E2F-6 from inhibiting the activity of the other E2F complexes. This strongly suggests that the observed repression is dependent upon an additional function(s) of the E2F-6 protein. DP-binding might still be required for this repression as an essential component of an actively repressive complex.

Clearly, the experiments described above do not allow us to address the mechanism of action of the endogenous E2F-6 species. In fact, it is still unclear whether or not E2F-6 functions as a repressor in vivo. It is possible to envisage alternate models for E2F-6 function. For
example, the DP proteins are known to be present in excess relative to the E2Fs in vivo and E2F-6 could exist to chaperone this pool of "unbound" DP. If correct, the "chaperone" model yields two clear predictions about the endogenous E2F-6 protein. (i) It should exist at sufficiently high levels in vivo to be able to bind to any DP protein that is not associated with E2F-1 through -5. (ii) E2F-6 should have a lower affinity/avidity for the DPs than any of the other E2F family members to ensure the efficient transfer of the DP from E2F-6 to E2F-1 through -5. Clearly, the "repressor" model of E2F-6 action also yields testable predictions. Regardless of whether or not its repressive properties are dependent or independent of its ability to bind to DNA, E2F-6 should have a similar or greater affinity for the DP proteins than the other E2Fs. Depending on its mechanism of action, it could also be present at either greater (if it acts through the sequestration of DP) or similar/ lower (if it regulates transcription in a sequence-specific manner) levels relative to the other E2Fs in vivo.

A direct comparison of the relative affinity/ avidity of the individual E2F family members for DP should help us to distinguish between these "chaperone" and "repressor" models, as will information about the relative abundance and DNA binding activity of the endogenous E2F-6 protein. Although we have yet to establish the true physiological role of E2F-6, its broad expression pattern suggests that it makes an important contribution to the regulation of E2F activity in vivo.
E. Materials and Methods

**cDNA identification and characterization**

GenBank, EMBL, and DDBJ databases were searched by using the TBLASTN program at the National Center for Biotechnology Information with a consensus E2F DNA-binding domain sequence derived from the primary amino acid sequence of the DNA-binding domains of E2F-1, -2, -3, -4, and -5. The identified E2F-6 human and mouse expressed sequence tags (ESTs) were obtained from Research Genetics (Huntsville, AL). A human fetal brain cDNA library (Stratagene) was screened with a 1.6 kbp EcoRI fragment of a human EST labelled with $[^\alpha-32P]$dCTP by random priming. Hybridization was performed at 42°C in 5x SSC/5x Denhardt's solution/30% formamide/0.5% SDS/dextran sulfate (50 μg/mL)/salmon sperm DNA (150 μg/mL). Filters were washed at 55°C for three twenty minute periods in 1x SSC/0.1% SDS. Positive clones were identified by autoradiography. Exonuclease III (New England Biolabs) digestion was used to generate nested deletions of both EST and cDNA clones, which were then sequenced with Sequenase 2.0 (United States Biochemical).

**Plasmid construction**

The human E2F-6 ORF was amplified by PCR with the primers 6.6
(GT TAGGATCCATGGCGCAGAAGTTACCAC) and 6.10
(CTCAGGATCCATCGTTGCTTACTTCAAG) and Vent polymerase. The PCR product was digested with BamHI and subcloned into pHACMV-neo-Bam to generate pCMV-HA-E2F-6. The plasmids pE2F4-CAT, pRSV-luciferase, pCMV-E2F-1, pCMV-E2F-4, pCMV-DP-1, pCMV-HA-DP-1, pCMV-HA-DP2, pCMV-RB, and pCMV-p107 have been described (Helin et al., 1993; Wu et al., 1995; Moberg et al., 1996). The 6x His-tagged E2F-6 vector was constructed by amplifying the E2F-6 ORF with primers 6.6 and 6.12
(CACTAAGCTTTATCAGTTGCTTACTTCCAGCA). The PCR product was digested with BamHI and EcoRI and subcloned into pQE30 (Qiagen, Chatsworth, CA).
Northern blot analysis

The cell line Northern blot [containing poly(A)+ RNA from 293 (adenocarcinoma), HeLa (cervical carcinoma), ML-1 (myeloid leukemia), T98G (neuroblastoma), MCF7 (breast cancer), or C33-A (cervical carcinoma)] or a human tissue blot [containing 2 μg (per lane) of poly(A)+ RNA isolated from the indicated tissues (CLONTECH)] was screened with probes corresponding to the full-length E2F-6 ORF or a 1,052-bp XmnI-SacI fragment from the 3’ untranslated region. These fragments were labelled with [α-32P]dCTP by using the Prime-It II kit (Stratagene). The blots were hybridized for 18 h at 65°C in 0.5 M sodium phosphate, pH 7.5/1 mM EDTA, pH 8.0/5% SDS/1% BSA and washed three times with 1x SSC/0.1% SDS at 65°C before autoradiography.

Polyclonal antibody production and Western blotting

The full-length 6x His-tagged human E2F-6 protein (amino acids 1-275) was expressed in bacteria, purified over a Ni2+-nitrilotriacetic acid-agarose resin (Qiagen), and used to immunize mice. Western blotting was performed as described by Moberg et al.

Transient transfection and in vitro assays

Cells were grown under standard conditions in DMEM supplemented with 10% fetal calf serum. Transient transfections were performed as described (Moberg et al., 1996). For the immunoprecipitation and gel retardation assays, transfections were conducted with 10 μg of each of the indicated plasmids plus pCMV-neo-Bam to give a total of 30 μg. Gel shift assays were carried out as described (Moberg et al., 1996) with unlabelled cell extracts normalized for total protein concentration. For immunoprecipitations, the cells were labelled with 250 μCi of [35S]-methionine Express labelling mix (NEN) in methionine-free medium (GIBCO/BRL) for 3.5 h. Immunoprecipitations were performed as described (Lees et al., 1993) with the following antibodies: I2CA5 [anti-haemaglutinin (HA) tag], KH20 [anti-E2F-1 (Helin et al., 1992)], LLF4-1 [anti-E2F-4 (Moberg et al., 1996)], sc-610x [anti-DP-1 (Santa Cruz Biotechnology)], sc-829x.
Precipitates were resolved on 10% SDS polyacrylamide gels and detected by fluorography. For transactivation assays, C33-A cells were transfected in duplicate with 4 μg of E2F4-CAT, 2 μg of RSV-luciferase (as an internal control for transfection efficiency), 14 μg of carrier DNA (pBIKKS+), and the indicated amounts of the pCMV-E2F expression vectors. These transfections were performed in the presence or absence of 3 μg of pCMV-HA-DP2. Within each experiment, the total concentration of CMV expression vector was kept constant by the addition of pCMV-neo-Bam. Chloramphenicol acetyltransferase (CAT) and luciferase assays were conducted as described in Lees et al (Lees et al., 1993).
F. Acknowledgements

We are grateful to members of the Lees laboratory for helpful discussion throughout the course of this study and to Patrick Humbert and Steve Bell for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health. J. A. L. is a recipient of the Professor Irwin Sizer and Helen Sizer Career Development Professorship.
G. References


FIP-1, a zinc-finger protein, interacts specifically with E2F-6.

B Fairchild, JM Trimarchi, and JA Lees

(Author's contribution: Identification of FIP-1 and mapping of the E2F-6 domain that binds FIP-1.)

A. Abstract

E2F transcription factors have been implicated in the control of G1/S-specific transcription and in the regulation of cellular proliferation. We recently reported the discovery of a novel E2F protein, E2F-6, that differs from other known E2F species in its lack of a transcriptional activation domain; E2F-6 is also the only E2F protein not known to be regulated by the pRB-family of transcriptional repressors. We now report the isolation of a novel, ubiquitously expressed protein interacting specifically with E2F-6 in yeast two-hybrid and co-immunoprecipitation assays. This novel protein, which we have named FIP-1, contains a putative C2-C2 zinc finger with homology to the DNA-binding zinc fingers of certain nuclear pore proteins. FIP-1 bears strong homology throughout its length to YAF-2, a YY1-binding protein of unknown function. FIP-1 binds specifically to the dimerization domain of E2F-6, and can associate with E2F-6 but not E2F-1, -2, -3, or -4 as determined by co-immunoprecipitation. Thus, FIP-1 may participate in E2F-6-specific functions or in their regulation.
B. Introduction

The E2F family of transcription factors bind DNA in a sequence-specific manner and regulate the induction of expression of a variety of genes at the G1/S transition. So far, eight genes have been identified that may contribute to this DNA-binding activity (Helin et al., 1992; Kaelin et al., 1992; Shan et al., 1992; Girling et al., 1993; Ivey-Hoyle et al., 1993; Lees et al., 1993; Beijersbergen et al., 1994; Ginsberg et al., 1994; Buck et al., 1995; Hijmans et al., 1995; Ormondroyd et al., 1995; Sardet et al., 1995; Wu et al., 1995a; Zhang and Chellappan, 1995; Rogers et al., 1996; Morkel et al., 1997; Cartwright et al., 1998; Gaubatz et al., 1998; Trimarchi et al., 1998). Their gene products can be divided into two related subgroups based on sequence homology. One subgroup, referred to as the DP proteins (DP-1 and -2) can heterodimerize with the members of the other subgroup, called the E2F proteins (E2F-1, -2, -3, -4, -5, and -6); this heterodimerization is required for efficient DNA binding by the complex (Bandara et al., 1993; Helin et al., 1993; Krek et al., 1993; Wu et al., 1995a; Trimarchi et al., 1998).

Although E2F activity is required in a general way for the regulation of G1/ S-specific transcription, the E2F proteins also have specific roles in the regulation of cellular proliferation and apoptosis. E2F-1, for example, is required for efficient apoptosis of mouse thymocytes (Field et al., 1996). E2F-3 is rate-limiting for proliferation of mouse embryonic fibroblasts in culture (P. Humbert, S. Dandapani, K. Rogers, and J. Lees, unpublished observations), and an antibody to E2F-3 inhibits S phase induction in a rat embryo fibroblast cell line (Leone et al., 1998). Both E2F-4 and E2F-5 have also been shown to have specific, required functions in mouse development (Lindeman et al., 1998), and P. Humbert, K. Rogers, S. Dandapani, and J. Lees, unpublished observations). In contrast, E2F-1, -2, -3, -4, and -5 can all promote the induction of S phase when overexpressed (Johnson et al., 1993; Beijersbergen et al., 1994; Qin et al., 1994; Shan and Lee, 1994; DeGregori et al., 1995; Schwarz et al., 1995; Lukas et al., 1996; Shan et al., 1996; DeGregori et al., 1997; Dirks et al., 1998), although E2F-5 may require the coexpression of a DP protein for this activity.
The E2F proteins are regulated at multiple levels, including transcription, localization, and protein stability. The best characterized mechanism regulating E2F activity is the association of E2F/DP complexes with members of the pRB family. pRB is the product of the retinoblastoma-susceptibility gene RB, which is mutated in 30% of human cancers (Weinberg, 1996). pRB and its homologs p107 and p130 are transcriptional repressors, and their binding to the transcriptional activation domain of E2F proteins not only inhibits activation but also represses basal transcription at E2F-regulated promoters (Weintraub et al., 1992; Bandara et al., 1993; Zamanian and La Thangue, 1993; Adnane et al., 1995; Bremner et al., 1995; Weintraub et al., 1995; Ferreira et al., 1998; Pierce et al., 1998). pRB, p107, and p130 bind preferentially to different E2F complexes: pRB can be found with E2F-1, -2, -3, or -4 (Moberg et al., 1996), while p107 associates only with E2F-4 (Moberg et al., 1996) and p130 binds to either E2F-4 or E2F-5 (Hijmans et al., 1995; Sardet et al., 1995). This association is believed to participate in the cell cycle-specific regulation of E2F target genes. At the approach of the G1/S transition, pRB becomes hyperphosphorylated by cyclin/CDK complexes (DeCaprio et al., 1989; Hinds et al., 1992; Ewen et al., 1993; Kato et al., 1993; Sherr, 1994; Lundberg and Weinberg, 1998), disrupting its association with E2F (Chellappan et al., 1991; Knudsen and Wang, 1997). pRB/E2F complexes are therefore believed to repress E2F transcriptional targets in early G1, whereas "free" E2F complexes activate transcription in late G1 and S phase. Some E2F complexes can be inactivated by cyclin A-mediated phosphorylation, which may contribute to the return of E2F activity to basal levels later in the cell cycle (Dynlacht et al., 1994; Krek et al., 1994; Xu et al., 1994; Kitagawa et al., 1995; Krek et al., 1995; Dynlacht et al., 1997).

The proper regulation of E2F activity is essential to the cell and the organism. Mice born with only one wild-type copy of the RB gene uniformly develop tumors, and mice completely deficient for pRB die in utero with defects in the regulation of S phase entry and apoptosis (Jacks et al., 1992); the increased tumorigenesis and inappropriate S phase entry and apoptosis are largely E2F dependent (Tsai et al., 1998; Yamasaki et al., 1998). Mice born without p107 and p130 die neonatally (Cobrinik et al., 1996). Overexpression of a form of E2F-1 bearing a deletion of the
cyclin A-binding domain in E2F-1 leads to a delay in S phase progression and an enhanced induction of apoptosis (Krek et al., 1995). Thus, inappropriate E2F activity can have a variety of deleterious effects. Furthermore, overexpression of dominant-negative forms of E2F-1, DP-1, or DP-2 can interfere with S phase entry (Dobrowolski et al., 1994; Wu et al., 1996; Fan and Bertino, 1997). Levels of E2F activity that are both sufficient and not excessive are therefore maintained by a variety of cellular mechanisms.

It is therefore of interest that, although E2F-6 contains DNA-binding and DP-dimerization domains with homology to the other E2F proteins, no transcriptional activation domain has been identified in E2F-6, nor does it interact with the pRB-family members. In fact, when overexpressed, E2F-6 represses the transcription of an artificial promoter bearing E2F sites (Morkel et al., 1997; Cartwright et al., 1998; Gaubatz et al., 1998; Trimarchi et al., 1998). While it has not been established that endogenous E2F-6 complexes repress E2F target genes, the lack of a domain mediating transcriptional activation or association with pRB-family members clearly suggests that, at a minimum, E2F-6 is likely to function in ways that are distinct from those of other E2F family members. The deregulation of E2F activity is associated with a deregulation of cellular proliferation and apoptosis; an understanding of the role(s) played by E2F-6 in these or other processes requires a knowledge of the in vivo activity of E2F-6, of the mechanisms regulating its activity and of the genes it regulates.

To date, analysis of E2F-6 DNA-binding activity and of the effects of E2F-6 on transcription have relied exclusively on overexpression of the protein. Although overexpression studies may suggest functions for a given gene, they are subject to a variety of limitations. One such limitation lies in the assumption that all proteins required for the function of that gene have been provided in the system. DP-1, for example, activates an E2F-responsive promoter very poorly when overexpressed by itself, but very efficiently when coexpressed with E2F-1 (Bandara et al., 1993; Helin et al., 1993; Krek et al., 1993). If E2F-6 complexes include another protein in vivo that has not been provided in the overexpression assays, these assays may not accurately recapitulate the normal functions of E2F-6. We have therefore executed a yeast two-hybrid screen.
to isolate proteins that can associate with E2F-6. We have identified a novel zinc-finger protein that we have named FIP-1 (E2F-6 Interacting Protein-1), and demonstrated that this protein interacts specifically with E2F-6.
C. Results

C1. Isolation of FIP-1

By yeast two-hybrid, we screened an activated human T cell library for proteins that could interact with an E2F-6 fusion protein. The E2F-6 fusion protein included most of the amino-terminus of this protein, as well as its putative DNA-binding and dimerization domains, fused in frame to the GAL4 DNA-binding domain; the T cell library transformants are fusions with the GAL4 transcriptional activation domain. We screened approximately $4 \times 10^5$ transformants, and identified four independent clones with the ability to interact specifically with the E2F-6 fusion protein but not with pRB or DP-1 fusion proteins. The strength of the protein-protein interactions in this two-hybrid system can be estimated by the relative abilities of the yeast clones to proliferate in the presence of increasing levels of an inhibitory drug (3-aminotriazole). By this criterion, the four E2F-6-interacting clones could be further subdivided into two clones that grew moderately well, indicating a moderate interaction, and two clones that grew very well, indicating a strong interaction. This distinction was also observed when the library plasmids were isolated from these clones and retransformed into the host yeast strain expressing the E2F-6 fusion protein, indicating that the differential growth properties of the yeast were indeed conferred by the library plasmid.

Both of the library plasmids conferring moderate growth in 3-aminotriazole encode fragments of DP-2. Both DP-1 and DP-2 have been shown to interact when overexpressed with E2F-6 as assayed by co-immunoprecipitation, and to cooperate with E2F-6 in a DNA-binding assay. The identification of DP-2 in the two-hybrid was therefore not an unexpected result. In contrast, no clones encoding DP-1 fragments were obtained. This suggests that the two-hybrid screen was not saturating. Indeed, E2F-6 does interact efficiently with DP-1 when the two are co-expressed in this two-hybrid system (data not shown), and DP-1 is known to be expressed in activated T cells (Moberg et al., 1996). Screening of additional transformants would therefore be expected to identify some plasmids encoding fragments of DP-1, and might reveal additional E2F-6-interacting proteins as well.
The library plasmids conferring robust growth in combination with E2F-6 in the two-hybrid encode fragments of the same novel protein, that we have named FIP-1. The library plasmid inserts include more than a kilobase of sequence, with a potential open reading frame of at least 228 amino acids (Figure 1). One of the FIP-1 clones begins at the twelfth amino acid in this potential open reading frame: (PKRQAKPAA . . .). The second clone begins more than one hundred nucleotides upstream. Two ATG codons in the correct frame, very close to each other, are present in this upstream region; both have similarity to the Kozak consensus sequence for initiating methionines, although it remains possible that we may not have isolated the entire open reading frame. It is perhaps worth noting that the region upstream of the potential initiating methionine(s) is extremely GC-rich, which would be consistent with proximity to a site of transcriptional initiation.

Although both library plasmid inserts encoding FIP-1 share a common 3′ end apparently resulting from the priming of an oligo-dT primer to a complementary tract of adenosine residues in the FIP-1 message, comparison of this sequence to known expressed sequence tags (ESTs) reveals that the FIP-1 message extends at least another kilobase in this direction (Figure 1). Multiple sequences derived from this 3′ region of the gene have been mapped to the short arm of chromosome 3 (Unigene entry Hs.56936; http://www.ncbi.nlm.nih.gov/UniGene/Hs.Home.html). These sequences are believed to reside near chromosomal marker D3S1261, which is located at 3p13-p14.1 (Hughes et al., 1994), and references therein). This region of chromosome 3 is the target of frequent deletion in a variety of human carcinomas (Yokoyama et al., 1992; Buchhagen et al., 1996; Corless et al., 1996; Hu et al., 1996; Iqbal et al., 1996; Leonard et al., 1996; Chu et al., 1998; Halbert et al., 1998; Rader et al., 1998; Venugopal et al., 1998).

The amino acid sequence of FIP-1 has multiple noteworthy features. FIP-1 has significant homology throughout its length to YAF-2, a protein identified in a two-hybrid screen with the transcription factor YY1 (Figure 2) (Kalenik et al., 1997). YAF-2 can bind directly to the zinc finger domain of YY1 in vitro and can counteract YY1-mediated repression in overexpression assays (Kalenik et al., 1997). The high degree of conservation between the amino-terminus of
Figure 1. (preceding page) Nucleotide sequence and predicted translation product of FIP-1. Nucleotides derived from the plasmids isolated in the two-hybrid screen are capitalized, whereas nucleotides that have only been inferred based on ESTs present in Genbank are in lower case. We have not confirmed the accuracy of these sequences; however, the sequences have been reported in at least four independent ESTs at each nucleotide position. The predicted translation product is depicted underneath the corresponding mRNA codons; the potential initiating methionines are depicted in bold-face. A solid line underlines the potential zinc finger, and a dotted line underlines the lysine/glutamic acid rich domain.

Figure 2. (following page) FIP-1 shares extensive homology with the YY1-associated factor YAF-2. The amino acid sequences of FIP-1 and YAF-2 were compared using the BESTFIT program of the Genetics Computer Group software package. Identities are depicted by a vertical line (1), and gaps are depicted with periods ( . ). The conserved putative zinc finger is depicted in bold-face.
Figure 2.

FIP-1  MTMGDKSPTPKRQAKPAADEGFWDC\textit{CSVCTFRNSAEAFKCSIC}DVRKGTST

YAF-2 .MGDKKSPTPKRHKPSDEGYWDC\textit{CSVCTFRNSAEAFKCMMC}DVRKGTST

FIP-1  RKPRINSQLVAQQVAAQTTPPKKEKEKEKVEKQDKEKPEKDEIKPSVTK

YAF-2  RKPRPSQLVAQQVTQFVPPTQSKKEKKDVKE...KEKSEKE.....TTSK

FIP-1  KNTNKTKPKSDLPPSAEANIQSANATKTSETNHSTRPRLKNDSTA

YAF-2  KNSHKKT.................................RPLKNDRSSA

FIP-1  QQLAVTVGNVTIITDFKEKTRSSSTSSSTVTSSAGSEQNNQSSSGESTDK

YAF-2  QHLEVTVGDLTVIITDFKEKTKSPPAS.....SAASADQHQQSGSSDNTER

FIP-1  GSSRSSTPKGDMSAVNDESF

YAF-2  GSRSSSPPGGEASSLNG...
YAF-2 and the proposed amino-terminus of FIP-1 is consistent with the possibility that one of the methionines highlighted in Figure 1 may indeed be the legitimate site of translational initiation.

FIP-1 shares with YAF-2 a C-X_2-C-X_{10}-C-X_2-C zinc finger motif in the amino-terminal region of the proteins (Figures 1 and 2). The zinc fingers of FIP-1 and YAF-2 are most highly homologous to nup153 and nup358 (Sukegawa and Blobel, 1993; McMorrow et al., 1994; Wu et al., 1995b; Yokoyama et al., 1995), both of which are believed to be associated with the nuclear pore. FIP-1 and YAF-2 also share a domain rich in lysine and glutamic acid. There are also differences between FIP-1 and YAF-2. FIP-1 bears a thirty-three amino acid insertion in the middle of the protein, absent in YAF-2. The carboxy-terminal region of FIP-1 is less well conserved to YAF-2 than most of the rest of the protein is (<40% identity vs. ~80%), and is rich in serine and threonine residues and poor in hydrophobic residues (5 of the last 52 residues). Interestingly, there are ESTs homologous to the FIP-1/ YAF-2 family from a variety of multicellular organisms including fruit flies and nematodes; the homology is particularly striking in the amino terminus (40/60 residues identical from D. melanogaster FIP-1/YAF-2 (GenBank entry AA264271, for example) to human FIP-1). Unfortunately, the functions of this gene in these organisms remain unknown.

C2. Expression pattern of FIP-1

We have investigated the pattern of expression of the endogenous FIP-1 gene by Northern blot analysis. The levels of FIP-1 mRNA expression vary among the human tissues assayed; levels were highest in samples derived from placenta, lung, and kidney, and were present at more moderate levels in muscle, liver, brain, pancreas, and heart (Figure 3). Despite the variations in expression levels, a 5 kb FIP-1 mRNA is expressed in all tissues examined; a second, slightly larger message is also detectable in the placental sample. A similar 5 kb message is also expressed in all cell lines we have examined to date (data not shown). E2F-6 mRNA has also been detected in these tissues and cell lines, consistent with the possibility that the two proteins may interact in vivo.
Figure 3.
Figure 3. (preceding page) FIP-1 is widely expressed *in vivo*. A PCR product corresponding to the entire putative open reading frame was radiolabelled and used to probe a human tissue northern blot (Clontech) as previously described. This probe does not cross to YAF-2 (data not shown); the YAF-2 message is also considerably smaller than the FIP-1 message (~2 kb vs. ~5 kb).
The existence of a ubiquitous 5 kb message indicates that we have not yet isolated a full-length FIP-1 cDNA. The 3 kb that we have not located could potentially result from additional sequences upstream and/or downstream of the 2 kb we have already identified. Theoretically, the additional sequences could also be interior, in whole or in part, to our current sequence as a result of alternative splicing; however, the lack of smaller detectable FIP-1 mRNA splice forms argues against this possibility.

C3. Identification of E2F-6 regions interacting with FIP-1

As a first step toward characterizing the interaction of FIP-1 with E2F-6, we have identified the domains of E2F-6 that mediate this interaction. We were interested, for example, to know if FIP-1 binds to a region of E2F-6 that is largely conserved in the other E2F proteins, such as the DNA binding domain, or to a more divergent region, such as the amino-terminus. We have therefore expressed, in the two-hybrid system, various domains of E2F-6 to determine which are necessary and which are sufficient for interaction with FIP-1 (Figure 4). Deletions from the carboxy-terminus of E2F-6 revealed that the domain believed to be required for interaction with the DP proteins, the "dimerization domain", including the "leucine zipper" and "marked box" regions, is also required for the interaction of E2F-6 with FIP-1. Deletions from the other direction indicate that neither the amino-terminal region nor the DNA-binding domain of E2F-6 is required for its interaction with FIP-1; indeed, the dimerization domain of E2F-6 is both necessary and sufficient for the interaction. This raises a number of specific questions. First, the similarity in the requirements of DP proteins and FIP-1 for binding to E2F-6 raises the possibility that their binding to E2F-6 may be mutually exclusive; this possibility is being actively investigated. Second, the dimerization domain of E2F-6 is somewhat conserved with other E2F proteins. We have therefore attempted to determine if FIP-1 can also interact with other members of the E2F family.
**Figure 4.** (preceding page) The dimerization domain of E2F-6 mediates its interaction with FIP-1. Fusion proteins expressing the indicated regions of E2F-6 fused to the GAL4 DNA-binding domain were expressed in yeast and their ability to interact with FIP-1 was assessed by yeast two-hybrid. The strength of the observed interaction, as inferred by resistance to 3-aminotriazole, is depicted in the right-hand column. "+++" indicates strong resistance to 3-aminotriazole, whereas "-" indicates susceptibility to the drug.
C4. Specific interaction of FIP-1 with E2F-6

We have subcloned the putative open reading frame of FIP-1 into a eukaryotic expression vector to assess its ability to interact with E2F proteins in mammalian cells; a haemagglutinin (HA) epitope tag was included at the amino-terminus of FIP-1. This construct expresses efficiently in transiently transfected C33-A cells, a human cervical carcinoma cell line, as determined by Western blotting and probing with an anti-HA antibody (data not shown).

The ability of FIP-1 to interact with E2F proteins was assayed by co-immunoprecipitation assays from $^{35}$S-labeled C33-A cells (Figure 5). Using an anti-HA antibody, FIP-1 is efficiently immunoprecipitated from extracts transfected with the construct expressing this protein; similarly, E2F-1, -2, -3, -4, and -6 can be specifically immunoprecipitated with the appropriate antibodies when overexpressed (Figure 5 and data not shown). E2F-6 can also be immunoprecipitated from cells co-expressing E2F-6 and FIP-1 and treated with an antibody to the HA tag on FIP-1, indicating that E2F-6 and FIP-1 are interacting in this assay and that E2F-6 is co-immunoprecipitated by virtue of its association with FIP-1 (Figure 5). Similarly, an antibody to E2F-6 immunoprecipitates both FIP-1 and E2F-6 when the two are coexpressed, although inefficiently: whereas nearly comparable amounts of E2F-6 are immunoprecipitated through an E2F-6 antibody or when co-immunoprecipitated with FIP-1, FIP-1 is clearly immunoprecipitated more efficiently through an anti-HA antibody than it is through the anti-E2F-6 antibody in the presence of co-expressed E2F-6. The reasons for this difference are currently under investigation.

Although FIP-1 interacts with E2F-6 in this assay, no interaction between FIP-1 and E2F-1, E2F-2, or E2F-4 can be detected. In cells overexpressing both E2F-1 and FIP-1, for example, treatment with an antibody to E2F-1 efficiently immunoprecipitated E2F-1, but not FIP-1, whereas an anti-HA antibody immunoprecipitated FIP-1, but not E2F-1. Similar results were observed for E2F-2, -3, and -4 (Figure 5 and data not shown). Thus, the capacity to interact with FIP-1 is not a general property of E2F proteins, but is specific to E2F-6.
Figure 5.

Transfected:

Antibody:

- HA
- E2F-1
- E2F-1
- E2F-1 + FIP-1
- E2F-1 + FIP-1
- E2F-2
- E2F-2 + FIP-1
- E2F-2 + FIP-1
- E2F-4
- E2F-4 + FIP-1
- E2F-4 + FIP-1
- E2F-6
- E2F-6 + FIP-1
- E2F-6 + FIP-1
Figure 5. (preceding page) FIP-1 associates specifically with E2F-6. Constructs expressing the proteins indicated at the top of the figure were transfected into C33A cells, which were subsequently labelled with \(^{35}\text{S}\)-methionine, lysed, and immunoprecipitated with an antibody to the epitope indicated at the bottom of the figure. The resulting proteins were resolved by SDS-PAGE. The lower arrow on each side of the figure indicates the position of FIP-1 protein, whereas the upper arrow on the left side indicates the approximate positions of E2F-1, -2, and -4, and the upper arrow on the right indicates the position of E2F-6.
D. Discussion

E2F activity participates in the regulation of cellular proliferation and apoptosis. The ubiquitous expression of E2F-6 suggests that it may play an important role in cellular regulation. Although E2F-6 has been shown, when overexpressed, to bind DP proteins and DNA and to repress E2F-mediated transcription, efforts to confirm these activities for endogenous E2F-6 protein have proven difficult. For example, endogenous E2F-6/DP proteins have not yet been reported, nor has E2F-6 yet been detected as a component of endogenous E2F DNA binding activity as assayed by standard E2F electrophoretic mobility shift assays (J. Trimarchi and J. Lees, unpublished data), perhaps consistent with its relatively low affinity for E2F binding sites in these assays when overexpressed.

The E2F literature is replete with examples of E2F-associated proteins that mediate or regulate the activity of these complexes. These range from the DP proteins, required for high affinity DNA binding and transcriptional activation by E2F-1 through -5, to the repressors pRB, p107, and p130, to cyclin A, which phosphorylates some E2F complexes, inhibiting their transcriptional activity. We have therefore executed an unbiased screen for E2F-6-associated proteins in an attempt to clarify the functions of endogenous E2F-6 and the mechanisms of the regulation of its activities. We have isolated a novel protein that, like E2F-6, is expressed in all tissues and cell lines examined, and is therefore a plausible candidate for a regulator or mediator of E2F-6 function. Our screen was not saturating; further screening has the potential to uncover additional E2F-6-interacting proteins.

This novel protein, FIP-1, bears significant homology to the YY1-associated protein YAF-2. Beyond its ability to bind YY1, there are few data to indicate the in vivo function(s) of YAF-2. Its message is upregulated upon differentiation of skeletal myotubes or cardiac myocytes (Kalenik et al., 1997). When overexpressed, YAF-2 can activate the expression of the YY1-repressible skeletal alpha-actin promoter (Kalenik et al., 1997). YAF-2 could therefore be hypothesized to counteract the effects of YY1 in some contexts. YAF-2 is not believed to encode a transcriptional
activation domain; rather, it has been suggested that YAF-2 may promote the specific degradation of YY1 upon muscle cell differentiation, although this remains highly speculative.

The putative zinc finger in FIP-1 (and YAF-2) is homologous to zinc fingers found in the nuclear pore proteins nup153 and nup358. These proteins have at least four zinc fingers each, all of the type W-x-C-x-C-x(3)-N-x(6)-C-x-x-C: the same pattern observed in the single zinc finger in FIP-1 and YAF-2 as well as in mdm2 and other proteins. Although mdm2 is known to interact with p53 through the amino-terminus of the former protein (Chen et al., 1993; Oliner et al., 1993) and has been reported to interact with E2F-1/DP-1 through that same region (Martin et al., 1995), no function has yet been ascribed to this mdm2 zinc finger, which is located near the middle of the protein. Both nup153 and nup358 have been reported to bind DNA in a zinc-dependent manner (Sukegawa and Blobel, 1993; Yokoyama et al., 1995); the zinc finger domain of nup153 is sufficient for this interaction. YAF-2 has been reported to have a moderate affinity for DNA cellulose (Kalenik et al., 1997). Given the low affinity of E2F-6/DP complexes for DNA in overexpression assays, it is tempting to speculate that FIP-1 may promote the interaction of E2F-6 complexes with DNA. Nevertheless, no data yet support that conclusion, and the zinc finger of FIP-1 may instead play a role in protein-protein interactions with E2F-6 or other proteins, or simply be required to maintain the overall structure of the protein. Importantly, although FIP-1 and YAF-2 share homology to nup153 and nup358 in their zinc finger, neither FIP-1 nor YAF-2 bears the FG repeats that are the hallmark of nuclear pore proteins.

Although the identification of FIP-1 as a specific E2F-6 binding protein is intriguing, it has not yet clarified in any way the in vivo functions of E2F-6. Many questions remain to be addressed. With what proteins is endogenous E2F-6 associated, and is FIP-1 among them? In what proportion? If endogenous E2F-6 is associated with FIP-1, many more questions arise. Does the presence of FIP-1 in the complex preclude the presence of DP proteins, or does the trimolecular E2F-6/ DP/ FIP-1 complex exist? Does FIP-1 promote DNA-binding by the complex? Does FIP-1 recruit additional proteins? Does it affect the stability of the complex, or its localization? Given that E2F-6 is the only known E2F that is not believed to be regulated by the
pRB family of proteins, it may have functions distinctly different from other E2F proteins, and will likely be regulated by different mechanisms. The identification of E2F-6-interacting proteins such as FIP-1, coupled with studies of endogenous E2F-6, should help to elucidate the particular roles of E2F-6 in the regulation of transcription and of cell biology.
E. Materials and Methods

Yeast two-hybrid

A fragment of E2F-6 whose amino-terminus was that reported in (ref) and carboxy-terminus was the end of the marked box (... PIDVYLCEVE) was cloned as a SacI/ SalI fragment into pc97 (Sardet et al., 1995) to generate an in-frame fusion with the DNA-binding domain of GAL4. This plasmid, called pc97-F6, was transformed into S. cerevisiae strain MaV103 (Sardet et al., 1995) using standard protocols. After establishing that the fusion protein neither interacts with the GAL4 activation domain nor transactivates on its own, this strain expressing pc97-F6 was transformed with a library derived from activated human T cells. The transformations were plated on SC-Leu-Trp plates. After two days, the transformations were replica-plated to SC-Leu-Trp-His plates supplemented with 10 mM 3-aminotriazole, a drug that selects for protein-protein interaction in this system (Sardet et al., 1995). After an additional two days, the transformations were again replica plated to SC-Leu-Trp-His plates, this time supplemented with 30 mM 3-aminotriazole. After two days, colonies were restreaked on SC-Leu-Trp-His +30 mM 3-aminotriazole plates. The library plasmid was isolated from the yeast and transformed into bacteria using standard protocols. Plasmids were re-transformed into MaV103 yeast expressing pc97-F6 to verify that 3-aminotriazole resistance was conferred by the library plasmid. Some sequencing was done manually using standard protocols, the remainder was done with an automated sequencer.

For mapping the FIP-1 binding region in E2F-6, all fragments were cloned as SacI/ SalI fragments into pc97 and transformed into MaV103 yeast expressing the library plasmids. The amino acid boundaries of the fragments are as follows: amino-terminus, HEKLPS... NVQYSMR; DNA-binding domain, KALKVK... KNHIRW; dimerization domain, IGSDL... . YLCEVE. Strength of interactions was inferred from degree of resistance to 3-aminotriazole in the two-hybrid assay.
Northern blot analysis

Northern blots were performed exactly as previously described (Trimarchi et al., 1998), except that the probe was derived from a PCR product corresponding to the predicted open reading frame of FIP-1 (MTMGDKK . . . VNDESF*). The blots were hybridized and washed exactly as previously described (Trimarchi et al., 1998).

Transient transfections and immunoprecipitations

Calcium phosphate were performed as previously described (Trimarchi et al., 1998). The eukaryotic expression vector for FIP-1, pCMV-HA-FIP-1, was generating by cloning a PCR product containing the predicted open reading frame of FIP-1 as a BamHI fragment into pHACMV-neo-Bam. Similarly, pCMV-E2F-6 was generated by cloning the BamHI fragment from pCMV-HA-E2F-6 into pCMV-neo-Bam. The plasmids pCMV-E2F-1, pCMV-E2F-2, pCMV-E2F-3, pCMV-E2F-4, and pCMV-DPl have all been previously described (Helin et al., 1993; Moberg et al., 1996).

Immunoprecipitations were performed as previously described with the following antibodies: 12CA5 [anti-haemagglutinin (HA) epitope tag], KH20 [anti-E2F-1 (Helin et al., 1992)], LLF2-5 [anti-E2F-2], LLF3-3 and LLF3-5 [anti-E2F-3], LLF4-1 [anti-E2F-4 (Moberg et al., 1996)], and LLF6-1 [anti-E2F-6]. The monoclonal antibodies LLF2-5, LLF3-3, and LLF3-5 were generated exactly as previously described (Moberg et al., 1996). For LLF6-1, mice were immunized with His-tagged E2F-6 protein as previously described (Trimarchi et al., 1998); splenocytes were isolated and monoclonal antibody-producing cell lines were generated as described (Moberg et al., 1996).
F. Acknowledgements

We are grateful to members of the Lees laboratory for helpful discussion throughout the course of this study and to Patrick Humbert and Steve Bell for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health. J. A. L. is a recipient of the Professor Irwin Sizer and Helen Sizer Career Development Professorship.
G. References


A. Cloning and characterization of novel E2F genes

A wealth of data indicate that the E2F transcription factors mediate the cell-cycle regulated transcription of a number of cellular genes at the G1/S transition, and that the proper regulation of these factors is necessary for accurate and timely passage through this transition and the subsequent S phase (reviewed in CHAPTER I). Consequences of their deregulation can also include apoptosis, transformation, and tumorigenesis. Both overexpression studies and genetic studies indicate that the six known E2F genes have some independent functions in addition to any shared ones. This evidence that the E2F genes are non-equivalent has provided much of the impetus for the research herein described. If the E2F genes have independent functions, only the identification and characterization of each of the E2F genes can, in this case, provide an accurate indication of the biological activities of E2F function in the cell. Similarly, the known E2F genes have been reported to differ in their regulation at the levels of mRNA abundance, protein stability, subcellular localization, or DNA-binding activity, again underlining the importance of studying each of these genes individually.

A1. What are the biological functions of E2F-6?

CHAPTERS III and IV detail the identification of E2F-6 and its early characterization. Although genetic studies with E2F-6 have not yet been carried out, overexpression studies suggest that E2F-6 may have biological functions that differ from those of the other E2F genes and suggests even more strongly that E2F-6 is likely to differ from other E2F genes in how it is regulated. For example, E2F-6 does not interact with pRB or p107 when overexpressed, and is therefore presumably independent of their regulatory effects; E2F-6 is the only E2F protein yet characterized for which this appears to be true. The lack of a transcriptional activation domain in
E2F-6 further suggests that the biological functions of this protein are likely to be different from those of other E2F family members. If E2F-6 functions as a transcriptional repressor as the overexpression data seem to suggest, then its activity is indeed different from that of the other E2F genes, conferring transcriptional repression in a manner independent of the pRB family. This repression could involve the recruitment of other transcriptional repressors and/or histone deacetylases. On the other hand, if E2F-6 can activate transcription, this function must be mediated by an accessory protein providing a transcriptional activation domain; the requirement for such a protein might provide an additional level of regulation of E2F-6 function.

The regulation of E2F-6 functions may vary among tissues or cell types, or among stages of differentiation or development. There are perhaps innumerable research avenues that could be explored to address this possibility. One approach would be to look for any variations in E2F-6 activity among cell types, tissues, or developmental stages. Although it is possible to monitor the effects of overexpressed E2F-6 on an E2F-dependent reporter construct in a variety of cell lines, and perhaps even to do so in cell lines that can be induced to differentiate, characterization of the endogenous E2F-6 activity is crucial. It is possible, for example, to study levels of E2F-6 protein in developing embryos by immunohistochemistry. Potentially more interesting would be a study of E2F-6 DNA binding activity, but conditions permitting the detection of endogenous E2F-6 DNA binding activity have not yet been reported. Whether this may point to a difference in the DNA binding specificity of endogenous E2F-6 proteins, a difference in required accessory proteins, or simply a difference in the assay conditions required has not yet been determined. The ability to identify and follow the DNA binding activity of the protein would provide a much more powerful tool to analyze the biochemistry of E2F-6 function, and may facilitate the study of its transcriptional activity as well as the study of its regulation in various cells and tissues.

Much as the other E2F complexes are known to oscillate between roles as transcriptional repressors and transcriptional activators, the transcriptional effects of E2F-6 could vary in different cellular contexts, depending perhaps on the presence or absence of particular accessory factors. An example of the identification and initial characterization of an E2F-6-interacting protein, FIP-1,
is described in CHAPTER IV. E2F-6-interacting proteins could conceivably provide or recruit
domains permitting E2F-6 either to activate or to repress transcription. Their presence or absence
in a particular cell or tissue could therefore determine the transcriptional effects of E2F-6 in that
specific environment. Alternatively, they could enhance or modify the DNA binding activity of
E2F-6. They could also regulate E2F-6 functions in other ways, such as at the level of protein
stability, or of subcellular localization. Characterizing the effects of E2F-6-interacting proteins on
the functions of E2F-6 may prove indispensable to a complete understanding of its biological
activities.

Determinaton of E2F-6 functions: a genetic approach

Although biochemical studies and the characterization of E2F-6-interacting proteins will
likely provide useful clues to the in vivo functions of E2F-6, a genetic study of E2F-6 will almost
certainly be required to fully identify and characterize those functions. By selectively eliminating
E2F-6 function in the mouse, for example, the contributions of E2F-6 to the regulation of E2F-
responsive genes can be accurately assessed, and the requirements for E2F-6 in the overall biology
of the organism can be determined. This approach has proven extremely useful in the assignment
of biological functions to other E2F family members: the specific requirements for E2F-4 in
hematopoiesis and for E2F-5 in choroid plexus function had not been indicated before they were
studied in mice lacking these genes, and the phenotypes associated with E2F-1 mutation strongly
support the overexpression studies indicating a pro-apoptotic function for this gene (reviewed in
CHAPTER I).

There is, of course, the possibility that the biological functions of E2F-6 may be shared
with other proteins; functional redundancy or functional compensation might therefore mask
functions of E2F-6 that might otherwise be revealed in an E2F-6-deficient animal. For other
genes, this difficulty has been partially circumvented by also mutagenizing other genes already
suspected to cooperate in the same biological function; this has proven fruitful in the intercrossing
of mice deficient for individual pRB family members, for example (Cobrinik et al., 1996; Lee et
In the case of E2F-6, it is much less clear with which other proteins it might be cooperating, principally because of the questions that remain regarding the in vivo function(s) of the protein. Despite any concerns that the study of E2F-6-deficient mice might not reveal every function of the gene, this approach almost certainly has the greatest chance to identify its in vivo function(s). This approach has often met with success in the study of other E2F family members and regulators; furthermore, no obvious alternative exists that can study effects in the whole animal and does not rely on overexpression. A genetic analysis may reveal particular tissues with a heightened requirement for E2F-6 function. Such information could serve as a focus for further research into E2F-6 activities. This approach will also provide cells lacking E2F-6 that could be used to analyze or screen for E2F-6 transcriptional targets, or used to determine if endogenous E2F-6 activity affects progression through the cell cycle.

A2. Additional E2F genes must exist

The total E2F DNA-binding activity in the cell is more than that which can be accounted for by the combined activities of E2F-1, -2, -3, -4, and -5. This is true both for the overall levels of E2F DNA-binding activity and for that subset that is associated with pRB. Although, when overexpressed with a DP subunit, E2F-6 can bind to an E2F DNA-binding site, endogenous E2F-6 does not appear to make a substantial contribution to E2F DNA-binding activity as determined by this standard gel shift assay. It cannot necessarily be inferred from this that E2F-6 is not, in fact, associated with E2F DNA sequences in vivo, merely that it does not appear to do so under the conditions of this assay. It can, however, be inferred that that one or more other E2F species must be responsible for the remaining E2F DNA-binding activity. Furthermore, since E2F-6 does not associate with pRB even when overexpressed, E2F-6 presumably does not account for the remaining pRB-associated E2F activity. Thus, there must exist at least one additional E2F species, and of these, at least one must be pRB-associated. The same arguments that previously impelled a search for E2F-6 therefore currently impel a search for the remaining E2F family members. The identification of E2F-6 has provided an additional piece of the E2F puzzle, and its obvious
differences from other E2F genes underscores the likelihood that its contributions to E2F function
are considerably different from those of other family members. Nevertheless, the final piece or
pieces of the puzzle have yet to be uncovered.
B. Toward an understanding of the mechanisms of E2F target specificity

As there is a need to identify and characterize E2F genes, there is a similar need to identify and characterize the targets of E2F cellular function. Because E2F complexes have been shown to have non-equivalent biological functions, much effort has recently been invested in determining how these different functions can be accomplished by a family of proteins with similar DNA-binding properties and (except E2F-6) similar abilities to transactivate or repress target genes in a manner dependent on the presence or absence of pRB family members. The differences in the biological functions of the pRB-associated E2F family members and their differential ability to affect individual E2F-responsive genes when overexpressed could, in principle, be mediated by a number of mechanisms. One possibility, that individual E2F complexes might vary in their DNA-binding specificities, was explored in Chapter II. The sequences of high-affinity DNA-binding sites for heterodimers of DP-1 with E2F-1, -2, -3, or -4 appear indistinguishable: they do not seem to determine which E2F complexes can bind to or regulate transcription through a particular DNA site. Another possibility is that specific E2F proteins may vary in their ability to interact with other promoter elements, leading to the specific stabilization of particular E2F complexes at individual promoters, depending on the other factors present at those promoters. E2F-1, -2, and -3, for example, have been proposed to interact with Sp-1 in vitro, whereas E2F-4 and -5 do not. E2F complexes obviously vary in their biological functions and in the transcriptional targets they can regulate, but the mechanism(s) of this specificity remain largely unresolved.

Chimeric E2F proteins as tools to decipher target specificity

Because specificity can be observed in overexpression assays, such as the specificity of E2F-1-induced apoptosis, or the specificity of E2F target gene activation, it should be possible to use these assays to identify the protein domain(s) conferring this specificity. Chimeric proteins have already been generated combining various regions of E2F-1 and E2F-4 or E2F-2 and E2F-4, and these hybrid proteins are competent to bind to DNA and activate transcription (Muller et al., 1997; Verona et al., 1997). Chimeric proteins combining regions of E2F-1 with regions of
another E2F, such as E2F-3 or E2F-4, could presumably be used in overexpression assays to
determine which region of E2F-1 mediates its proapoptotic function. If, for example, a version of
E2F-3 whose DNA-binding domain has been replaced with that of E2F-1 can drive apoptosis
when overexpressed, then the manner in which E2F-1 associates with DNA could be important for
this effect, or perhaps E2F-1 associates through this domain with an additional protein affecting
promoter selection or transcriptional regulation. Regardless of which domain proves to mediate
this function, increasingly smaller regions of E2F-1 and E2F-3 could be interchanged until the key
residues for the promotion of apoptosis are identified, which would hopefully lead to the rapid
identification of the relevant biochemical activity (protein-protein binding, protein-DNA binding,
etc.).

Much as overexpression assays could be used to determine the function(s) of E2F-1
involved in the promotion of apoptosis, they could also be used to identify elements conferring
specificity of target gene activation. For example, E2F-1 has been reported to be more efficient
than E2F-3 in the induction of ribonucleotide reductase 2 (RR2) mRNA levels, whereas E2F-3 is
the more efficient at inducing CDK2 (DeGregori et al., 1997). Overexpression of chimeric E2F-1/
E2F-3 proteins should identify the residues of these proteins conferring this specificity. Similarly,
the CDK2 promoter could be altered to more closely resemble the RR2 promoter (location of E2F
site, sequence and orientation of E2F site, presence or absence of binding sites for other
transcription factors) to identify promoter elements conferring this specificity. The specificity of
target gene activation has only been reported to occur with respect to the expression of endogenous
target genes. The ability of these overexpression assays to maintain specificity in
extrachromosomal constructs has therefore not yet been determined, and the identification of the
promoter elements conferring specificity may require the modification of the endogenous loci.
This, although more laborious than studies involving reporter plasmids, should nevertheless be
feasible for a limited number of mutations, and might be very productive when used in conjunction
with chimeric protein experiments.
**Coupling E2F biology to E2F target genes**

Understanding how E2F regulates transcriptional targets and understanding how these promoters are targeted specifically by individual E2F complexes are of interest primarily insofar as they relate to cellular and organismal biology. Overexpression of a version of E2F-1 lacking the residues required for its pro-apoptotic function, for example, could potentially be used to identify the transcriptional target genes mediating this effect. The overexpression of wild-type E2F-1 induces p19-ARF mRNA levels more efficiently than the overexpression of E2F-2, -3 or -4 does (DeGregori et al., 1997), and p19-ARF has been proposed to mediate at least some of the pro-apoptotic effects of E2F-1 (Bates et al., 1998). It would be of interest to determine if the residues required for these pro-apoptotic effects also mediate the specificity of p19-ARF induction by E2F-1. Other potential E2F-1 target genes could also be assayed for their dependence on any pro-apoptotic domain of E2F-1. It should even be possible to screen for novel genes whose expression can be induced by wild-type E2F-1 but not by a version lacking the pro-apoptotic residues; such genes would be additional candidates for mediators of the pro-apoptotic functions of E2F-1.

Although other E2F genes have also been shown to have specific functions, these have generally been in loss-of-function assays (e.g. deletion, or antibody-mediated inactivation), and may therefore require other approaches to match deregulated target genes to biological consequences. E2F-3-deficient MEFs, for example, proliferate more slowly than wild-type cells (P. Humbert, K. Rogers, S. Dandapani, and J. Lees, unpublished data). This is presumably mediated by inappropriate levels of one or more E2F-3 target genes, resulting either from a lack of E2F-3/DP-mediated transcriptional activation or from a lack of pRB/E2F-3/DP-mediated transcriptional repression; the overall level of expression could be inappropriate, or merely its timing during the cell cycle. A careful study of the expression of E2F target genes in these cells would therefore be predicted to identify one or more genes whose timing or levels of expression differ from wild-type expression patterns; some subset of these E2F-3 target genes are presumably involved in the reduced proliferation rate of E2F-3 deficient MEFs. In contrast, E2F-4-deficient
MEFs have an approximately wild-type proliferation rate (P. Humbert, K. Rogers, S. Dandapani, and J. Lees, unpublished data). Any E2F target genes deregulated in both E2F-3-deficient MEFs and in E2F-4-deficient MEFs are therefore less likely to be involved in the proliferative effects of E2F-3 than are target genes only deregulated in E2F-3 deficient MEFs. In contrast, loss of E2F-4 function disrupts hematopoiesis (P. Humbert, K. Rogers, S. Dandapani, and J. Lees, unpublished data), and loss of E2F-5 deregulates cerebrospinal fluid production in the choroid plexus (Lindeman et al., 1998); an analysis of E2F target gene expression in these tissues of mutant mice may enhance the current understanding of these functions of E2F-4 and -5. Indeed, cells from these mice could also be used in screens to identify novel E2F transcriptional targets that may be specifically regulated by these E2F complexes.
C. Broadening our understanding of E2F: tumor genetics and genetic screens

C1. E2F and cancer genetics

**E2F-1 as a tumor suppressor**

Of particular interest to cancer biologists are the roles played by the E2F genes in tumorigenesis. A loss-of-function mutation of mouse E2F-1 is associated with late-onset tumors of the reproductive tract (Yamasaki et al., 1996). The causes of this tumorigenesis are unknown. It could result from the failure to express an E2F-1-responsive oncogene, or from a derepression of a pRB/E2F-1/DP-regulated tumor suppressor. Whether this tumorigenesis results from a cell autonomous defect in these mice has not been reported, nor has it been reported if the tumors that also arise in mice heterozygous for an E2F-1 mutation display a loss-of-heterozygosity at the locus. The relevance of this tumorigenesis to human cancer, moreover, has not yet been established; no tumor-specific alterations in the human E2F-1 gene have been reported. This is perhaps surprising, unless the defect is not cell autonomous, in which case its specific detection in tumors would not be expected.

**pRB inactivation, tumorigenesis, and E2F**

Inactivation of pRB, in contrast, is known to be an extremely frequent event in human cancers, either through loss-of-function mutations in the RB gene, mutations in the cyclins, CDKs, and CDK inhibitors that modulate pRB phosphorylation, or expression of viral oncoproteins directly inactivating pRB, p107, and p130. Understanding the effects of E2F deregulation in this context could answer many important questions. First, such an understanding would address the degree to which E2F deregulation is a (or perhaps the) critical functional target in pRB-suppressed tumorigenesis. Second, it would determine whether derepression of pRB-regulated transcriptional targets is sufficient for tumorigenesis, or if there is a further requirement for E2F-mediated transcriptional activation of these targets. Third, it would identify the subset of E2F genes relevant
to tumorigenesis. Fourth, it would also, ideally, identify the transcriptional targets of this subset that promote tumor formation and growth.

This work has already begun with studies done on E2F-1 inactivation in the background of SV40 expression or RB heterozygosity in the mouse. In the context of SV40-induced tumors of the choroid plexus, E2F-1 mutation simultaneously reduces apoptosis five-fold, presumably accelerating tumor growth, and reduces cell cycle progression, leading to an unchanging overall rate of tumor growth (Pan et al., 1998). In mice heterozygous for an RB loss-of-function mutation, deletion of E2F-1 prevents tumorigenesis in the thyroid (0% penetrance versus 53% in the presence of E2F-1) and reduces its occurrence in the pituitary (62% penetrance versus 95%) (Tsai et al., 1998). Thus, in at least one tissue the induction of tumorigenesis by the loss of pRB requires the deregulation of E2F-1, and in another tissue the efficiency of tumorigenesis requires E2F-1. Deletion of E2F-1 in a background of RB heterozygosity also extends the mouse life span, presumably as a direct result of decreased tumorigenesis (Tsai et al., 1998); the decreased penetrances of the tumor phenotypes are therefore more striking, given that the tumors have had additional time to arise. These results also indicate that the transcriptional activation function of E2F-1 is essential for the efficient tumorigenesis of RB heterozygotes and for efficient tumor-associated apoptosis and cell proliferation in the SV40 transgenic mice. Interestingly, tumorigenesis in the thyroids of RB-heterozygous mice may also be dependent on the dosage of E2F-1: tumor penetrance in this tissue is reduced from 53% to 6% by heterozygous mutation of the E2F-1 gene (Tsai et al., 1998), arguing that only maximal transcriptional activation by E2F-1 is sufficient for efficient tumorigenesis in this tissue.

While these studies clearly argue that E2F activity is a critical target for tumorigenesis induced by loss of pRB, other important questions remain unresolved. One is the degree of involvement of other E2F genes in this pathway. The effects of E2F-1 gene dosage in thyroid tumorigenesis, for example, could be explained by a requirement for high levels of an E2F-1-specific target gene, or by a requirement for high levels of a function shared among the E2F proteins, such as the regulation of S phase entry. Studies on the effects of mutations in other E2F
genes in the background of pRB inactivation may help to address this question. If mutation of E2F-2, -3, or -4 phenocopies a mutation of E2F-1, that may argue that a function shared among the E2F genes is required for efficient tumorigenesis, and that the mutation of any of these genes merely drops the overall E2F levels below a threshold requisite for tumorigenesis. The formal possibility that a different function for each E2F gene, required in for tumorigenesis in the same tissues and to the same degree as for E2F-1, would remain, but would seem a much less likely explanation. If, on the other hand, mutation of the other E2F genes does not affect tumorigenesis, or affects tumorigenesis in different tissues, this would indicate a very specific requirement for the deregulation of particular E2F complexes in tumorigenesis.

Identifying the E2F genes that promote tumorigenesis will focus future research

The identification of the E2F proteins relevant to tumorigenesis upon pRB inactivation is important for at least two reasons. (i) If, for example, E2F-1 has a role in the tumorigenesis of some tissues that is not shared with other E2F genes, then the search for E2F transcriptional targets relevant to tumorigenesis can be narrowed to E2F-1-specific targets in these tissues. (ii) Research can be focused on this branch or branches of the E2F pathway in an attempt to identify opportunities for clinical intervention. Given that the E2F proteins are regulated in different ways, it might be possible to selectively inactivate those complexes required for tumorigenesis. A generalized inhibition of E2F complexes may prove to have wide ranging effects in vivo: the introduction of double-stranded DNA bearing E2F DNA binding sites into smooth muscle cells, for example, inhibits their proliferation in rats (Morishita et al., 1995). In contrast, it has been suggested that an in vitro selection scheme that has isolated an RNA ligand capable of inactivating E2F proteins 1-5 and inhibiting G1/S-progression could be modified to identify ligands selectively inhibiting individual E2F complexes (Ishizaki et al., 1996). If so, such a ligand might someday be used in cancer patients to inactivate only those E2F complexes that contribute to tumor growth, while minimizing any side effects associated with the additional inactivation of other E2F complexes.
C2. Genetic screens could uncover additional E2F functions or regulators

"Reverse genetics", in which an animal with an engineered mutation is studied to further the understanding of a particular gene, can be invaluable in assessing the functions of individual E2F genes in biological pathways. Nevertheless, these studies have intrinsic limitations. One such limitation is that although these studies can reveal biological functions for which the gene in question is essential, they do not reveal functions for which the activity of the gene is simply sufficient, but not essential. E1A, for example, is known to promote cellular proliferation both through regions of the protein that inactivate pRB family members and through regions that bind p300 and CBP; either region of the protein can be sufficient for this effect (Zerler et al., 1987; Bellett et al., 1989; Howe et al., 1990). It is possible, by extension, that in some cellular contexts E2F transcription might be sufficient but not required to promote entry into the cell cycle. Mouse studies in which E2F genes have been inactivated would not reveal such a function.

The identification of biological activities that the E2F pathway may share with other signalling pathways may require other approaches, such as genetic screens in Drosophila or other genetically tractable organisms. There may be differences between the functions of E2F in Drosophila and in mammals; any function suggested by studies in other organisms would need to be confirmed in mammals. Nevertheless, overexpression studies and mutational studies indicate that in both Drosophila and mammals E2F activity is required for the G1/S-specific transcription of genes involved in DNA synthesis and cell cycle control, that E2F activity promotes cell cycle progression and that an inhibitor of E2F inhibits cell proliferation, and that the overexpression of an E2F gene can induce apoptosis (reviewed in CHAPTER I). Thus, many functions of E2F have been conserved among these species, and it seems reasonable to expect that any additional functions attributed to the pathway might be similarly conserved. Furthermore, Drosophila has recently been shown to have at least two E2F genes that vary in their transcriptional activity (Sawado et al., 1998), and E2F-responsive genes in Drosophila have been reported to vary in the
extents to which their responsiveness results from E2F-mediated repression or E2F-mediated activation, suggesting additional similarities to mammalian E2F regulation.

Genetic screens could reveal much about the functions and regulation of E2F. It is possible, for example, to screen for mutations that are synthetically lethal with weakly hypomorphic mutations in the E2F pathway, or for mutations accelerating the lethality of a dE2F or dDP null mutation. Such a screen could uncover new biological functions of E2F in the regulation of growth, proliferation, apoptosis, differentiation, or other aspects of development. At the same time, it would reveal genes that regulate E2F function (such as pRB family members), genes that participate in E2F function (such as additional E2F/DP family members), critical targets of E2F function (such as cyclin E), or genes in other signalling pathways that cooperate with E2F to induce these biological effects. Other potential genetic screens to identify modifiers of E2F function include screening for mutations affecting the expression pattern of E2F-responsive genes, a technique that has already successfully identified mutations both in dDP and in other complementation groups (Royzman et al., 1997); the proper development of the Drosophila eye is very sensitive to the gene dosage of cell cycle regulators, including dE2F and RBF (Du et al., 1996; Secombe et al., 1998) and could also be a basis for screening for mutations interacting genetically with E2F.

C3. Genetics, molecular biology and biochemistry in the study of E2F

The importance of E2F regulation was first suggested by the genetics of E1A, the genetics of RB, and the realization that E1A releases E2F from pRB. Biochemical and overexpression studies of E2F function have revealed much about its roles in transcriptional activation and repression, and have suggested functions in the regulation of cell biology, some of which have been confirmed genetically. These biochemical and overexpression studies may yet give us more insight into the ways E2F is regulated, and into the ways E2F regulates transcription and cell biology. Molecular biology will likely reveal the missing participants in E2F activity. Ultimately, however, the genetics will confirm or disprove the importance of E2F in tumorigenesis, apoptosis,
and other aspects of cell biology, and the genetics may prove the most efficient means to identify additional regulators of E2F function, additional mediators of E2F function, and additional E2F functions in transcriptional control, cell biology, and organismal biology.
D. References


ACKNOWLEDGEMENTS

Those who have helped me over the course of this work are many, very many, and I thank you each and all. Jackie deserves a special note of thanks for accepting me into her new lab and encouraging me to learn what it is to be a scientist. Special thanks, too, should go to all of the members of the Lees lab who gave their support in the times of scientific bounty and of drought. I thank Joel Pomerantz who gave me insights into "graduate studentdom" that I am still unraveling and appreciating to this day. I would also like to thank all of the people of the fifth floor who made lab a more interesting and enjoyable place to be. I doubt I would have made it this far without the constant support of my wife Susan, to whom I am particularly grateful. All of you I have listed here, and many more, had a big part in this work. It would not have happened without you. Thank you.