

A Novel E2F-like Protein Involved in Transcriptional Activation of Cyst Wall Protein Genes in *Giardia lamblia**^[S]

Received for publication, July 7, 2011, and in revised form, August 9, 2011. Published, JBC Papers in Press, August 11, 2011, DOI 10.1074/jbc.M111.280206

Li-Hsin Su, Yu-Jiao Pan, Yu-Chang Huang, Chao-Cheng Cho, Chia-Wei Chen, Shao-Wei Huang, Sheng-Fung Chuang, and Chin-Hung Sun¹

From the Department of Parasitology, College of Medicine, National Taiwan University, Taipei 100, Taiwan

Giardia lamblia differentiates into resistant walled cysts for survival outside the host and transmission. During encystation, synthesis of cyst wall proteins is coordinately induced. The E2F family of transcription factors in higher eukaryotes is involved in cell cycle progression and cell differentiation. We asked whether *Giardia* has E2F-like genes and whether they influence gene expression during *Giardia* encystation. Blast searches of the *Giardia* genome database identified one gene (*e2f1*) encoding a putative E2F protein with two putative DNA-binding domains. We found that the *e2f1* gene expression levels increased significantly during encystation. Epitope-tagged E2F1 was found to localize to nuclei. Recombinant E2F1 specifically bound to the thymidine kinase and *cwp1–3* gene promoters. E2F1 contains several key residues for DNA binding, and mutation analysis revealed that its binding sequence is similar to those of the known E2F family proteins. The E2F1-binding sequences were positive *cis*-acting elements of the thymidine kinase and *cwp1* promoters. We also found that E2F1 transactivated the thymidine kinase and *cwp1* promoters through its binding sequences *in vivo*. Interestingly, E2F1 overexpression resulted in a significant increase of the levels of CWP1 protein, *cwp1–3* gene mRNA, and cyst formation. We also found E2F1 can interact with Myb2, a transcription factor that coordinate up-regulates the *cwp1–3* genes during encystation. Our results suggest that E2F family has been conserved during evolution and that E2F1 is an important transcription factor in regulation of the *Giardia* *cwp* genes, which are key to *Giardia* differentiation into cysts.

Giardia lamblia is an intestinal pathogen that causes waterborne diarrheal disease worldwide (1–3). Its infection contributes greatly to malnutrition and malabsorption leading to delayed child development (4). Parasitic trophozoites attach to the small intestine mucosa and differentiate into infective cysts when carried downstream to the lower intestine (5, 6). Transmission of giardiasis occurs through ingestion of infective cysts from contaminated water or food (2). *G. lamblia* is a valuable

model for basic studies as its life cycle can be reproduced *in vitro* by mimicking the intestinal environment (5, 6).

In addition to its important role as a pathogen, *G. lamblia* also raises great biological interest for understanding the progress of eukaryotic evolution. *G. lamblia* lives independently as a single-celled eukaryote, and it is classified as a protist. Phylogenetic trees obtained from the small subunit ribosomal RNA and translation elongation factors revealed *G. lamblia* as an early diverging eukaryote (7–9), but some phylogenetic trees support *G. lamblia* as a position that diverged nearly simultaneously with the opisthokonts and plants (10). *G. lamblia* has many special features that are biologically different from those of higher eukaryotes (2, 11). It lacks clear homologs to many cellular components of DNA synthesis, transcription, and RNA processing, supporting the idea that the missing components are too different to be recognized or they are nonessential (11). Many aspects of giardial transcription are unusual. Known giardial transcription factors have diverged at a higher rate than those of crown group eukaryotes (12). Only four of the 12 general transcription initiation factors have giardial homologs (11, 12). Giardial TATA-binding protein is highly divergent with respect to archaeal and higher eukaryotic TATA-binding proteins (12). Giardial RNA polymerase II has no regular heptad repeats in the C-terminal domain, and transcription by RNA polymerase II is highly resistant to α -amanitin (12, 13). The giardial promoter regulatory mechanism may be unusual because very short 5'-flanking regions (<65 bp) with no consensus TATA boxes or typical *cis*-acting elements are sufficient for expression of many genes (14–20). Interestingly, AT-rich sequences have been found around the transcription start sites of many genes (14–23). They are essential for promoter activity and play a predominant role in determining the positions of the transcription start sites, functionally similar to the initiator (Inr)² elements in higher eukaryotes (18–20, 24).

G. lamblia cysts can survive in hostile environments, such as fresh water and gastric acid, during infection as they have a protective wall composed of proteins and polysaccharides (5, 6). Three known cyst wall proteins are highly synthesized in a concerted manner during differentiation into cysts (encystation) (15, 16, 23). However, little is known of the detailed mechanisms governing their synthesis. During encystation, three *cwp* genes are coordinately induced (15, 16, 23), suggesting the importance of gene regulation at transcriptional and/or post-transcriptional level.

* This work was supported by the National Science Council Grants NSC 98-2320-B-002-018-MY2, NSC 99-2320-B-002-017-MY3, and NSC 100-2325-B-002-039, National Health Research Institutes Grants NHRI-EX98-9510NC and NHRI-EX99-9510NC in Taiwan, and in part by the Department of Medical Research in National Taiwan University Hospital and the Aim for the Top University Program of National Taiwan University.

^[S] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Fig. S1.

¹ To whom correspondence should be addressed. Tel.: 886-2-23123456-88262; Fax: 886-2-23915294; E-mail: chinhsun@ntu.edu.tw.

² The abbreviations used are: Inr, initiator; RACE, rapid amplification of cDNA ends.

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

Few transcription factors regulating *cwp* gene expression have been identified (21, 24, 25). A Myb family transcription factor (Myb2) is encystation-induced and is involved in coordinate up-regulation of *cwp1–3* genes and its own gene by binding to specific sequences (21, 26). Two GARP family proteins are involved in transcriptional regulation of many different genes, including the encystation-induced *cwp1* gene by binding to specific sequences (25). An AT-rich interaction domain family transcription factor can bind to specific AT-rich Inr sequences and function as an important transactivator in the regulation of the *cwp1* gene (24). WRKY can bind to specific sequences in the *cwp1*, *cwp2*, and *myb2* promoters and up-regulate expression of these genes (27). It has also been suggested that the *cwp* mRNA stability is regulated by factors of an incomplete nonsense-mediated mRNA decay pathway (28, 29).

E2F was named because it is an adenovirus E2 gene promoter binding factor (30). E2F proteins play important roles in regulating cell cycle progression, cell proliferation, and differentiation (31–36). They may help enter S phase by activating genes required for DNA synthesis and S phase entry, including thymidine kinase, DNA polymerase, and histones (31–33, 35). E2Fs are regulated by retinoblastoma family proteins, including pRB, p107, and p130 (37–41). These proteins may bind E2F transactivation domain to inhibit the transactivation function of E2Fs (37–42).

E2F proteins contain one or two ~70-amino acid DNA-binding domains with an RRYD motif (43). Eight E2Fs have been found in humans to date (44). Human E2F1–6 have one DNA-binding domain. DP is another family of transcription factors that can bind DNA and activate transcription (44). The DNA-binding domain of DP has significant similarity to that of E2F (44). E2F can form a heterodimer with DP to recognize a sequence TTT(c/g)GCGC(c/g) or TTTCCCGCC and to transactivate (E2F1–3) or repress (E2F4–6) the target gene promoters (42, 45–52). Human E2F7 and -8 have two DNA-binding domains. They can bind the consensus E2F-binding sequence in a DP-independent manner, but they function as transcriptional repressors to regulate target genes (43, 53, 54). Members of the E2F and DP family have been found in flies, worms, plants, and mammals (42). E2Fs and DPs with one DNA-binding domain have been identified in plants with similar function and similar binding sequence. E2Fs with two DNA-binding domains have also been found in plants with similar binding sequence, but they function as transcriptional repressors (E2L1–3) (55).

Because E2F proteins play critical roles in cell cycle progression and cell differentiation of many eukaryotes, we asked whether *Giardia* has E2F family proteins and whether they influence gene expression during *Giardia* differentiation into dormant cysts. We searched the *Giardia* genome database for genes encoding E2F-like proteins and identified one giardial E2F homolog (E2F1). We found that E2F1 bound to specific sequences and functioned as a transactivator of the thymidine kinase gene. We also found that E2F1 can function as a transcriptional activator of the *cwp* genes and interact with an important encystation-specific Myb2 transcription factor. Our results suggest that E2F1 may be an important transcription

factor regulating differentiation in the primitive eukaryote *G. lamblia*.

EXPERIMENTAL PROCEDURES

***G. lamblia* Culture**—Trophozoites of *G. lamblia* WB (ATCC 30957), clone C6, were cultured in modified TYI-S33 medium (56). Encystation was performed as described previously (23). Briefly, trophozoites that were grown to late log phase in growth medium were harvested and encysted for 24 h in TYI-S-33 medium containing 12.5 mg/ml bovine bile at pH 7.8 at a beginning density of 5×10^5 cells/ml.

Cyst Count—Cyst count was performed on stationary phase cultures ($\sim 2 \times 10^6$ cells/ml) during vegetative growth as described previously (57), and the data are shown in Fig. 8D. Cells were subcultured in growth medium with a suitable selection of drugs at an initial density of 1×10^6 cells/ml. Cells seeded at this density became confluent within 24 h. Confluent cultures were maintained for an additional 8 h to ensure that the cultures were in stationary phase (at a density of $\sim 2 \times 10^6$ cells/ml). Cyst count was performed on these stationary phase cultures. Cultures were chilled, and cells were washed twice in double-distilled water at 4 °C, and trophozoites were lysed by incubation in double-distilled water overnight at 4 °C. Cysts were washed three times in double-distilled water at 4 °C. Water-resistant cysts were counted in a hemocytometer chamber. Cyst count was also performed on 24-h encysting cultures.

Isolation and Analysis of the *e2f1* Gene—The *G. lamblia* genome database (11, 58) was searched with the amino acid sequences of the human E2F1 (GenBank™ accession numbers U47675 and U47646) using the BLAST program (59). This search detected one putative homolog for E2F, which was named as E2F1 (GenBank™ accession number XM_001705535, open reading frame 23756 in the *G. lamblia* genome database). The E2F1 coding region with 350 nucleotides of 5'-flanking regions was cloned, and the nucleotide sequence was determined. The *e2f1* gene sequence in the database was correct. To isolate the cDNA of the *e2f1* gene, we performed RT-PCR with *e2f1*-specific primers using total RNA from *G. lamblia*. For RT-PCR, 5 μg of DNase-treated total RNA from vegetative and 24-h encysting cells was mixed with oligo(dT)_{12–18} and random hexamers and Superscript II RNase H-reverse transcriptase (Invitrogen). Synthesized cDNA was used as a template in subsequent PCR with primers E2F1F (CACCATGCAGCTCGCAGGCCAA) and E2F1R (GTGT-GCTCCGTTTGCGTG). Genomic and RT-PCR products were cloned into pGEM-T easy vector (Promega) and sequenced (Applied Biosystems).

RNA Extraction, RT-PCR, and Quantitative Real Time PCR Analysis—Total RNA was extracted from *G. lamblia* cell line at the differentiation stages indicated in the legends to Figs. 2 and 8 using TRIzol reagent (Invitrogen). For RT-PCR, 5 μg of DNase-treated total RNA was mixed with oligo(dT)_{12–18} and random hexamers and Superscript II RNase H⁻ reverse transcriptase (Invitrogen). Synthesized cDNA was used as a template in subsequent PCR. Semi-quantitative RT-PCR analysis of *e2f1* (XM_001705535, open reading frame 23756), *e2f1-ha*, *cwp1* (U09330, open reading frame 5638), *cwp2* (U28965, open reading frame 5435), *cwp3* (AY061927, open reading frame

2421), *myb2* (AY082882, open reading frame 8722), thymidine kinase (XP_001705197.1, open reading frame 8364), *ran* (U02589, open reading frame 15869), and 18 S ribosomal RNA (M54878, open reading frame r0019) gene expression was performed using the following primers: E2F1F and E2F1R, E2F1HAF (TCCCACCCCCTCCATCAC) and HAR (AGCGT-AATCTGGAACATCGTATGGGTA); *cwp1F* (ATGATGCT-CGCTCTCCTT) and *cwp1R* (TCAAGGCGGGGTGAGGCA); *cwp2F* (ATGATCGCAGCCCTTGTCTA) and *cwp2R* (CCT-TCTGCGGACAATAGGCTT); *cwp3F* (ATGTTTTCTCTG-CCTTCTTCT) and *cwp3R* (TCTGTAGTAGGGCGGCTGTA); *myb2F* (ATGTTACCGGTACCTTCTCAGC) and *myb2R* (GGGTAGCTTCTCACGGGGAAG); *tkF* (ATGAACCTCCCT-TACGCTC) and *tkR* (CAGGTCTTCGTTGATCCC); *ranF* (ATGTCTGACCCAATCAGC) and *ranR* (TCAATCATCGT-CGGGAAG); 18SrealF (AAGACCGCCTCTGTCAATCAA) and 18SrealR (GTTTACGGCCGGGAATACG), respectively. For quantitative real time PCR, SYBR Green PCR master mixture was used (Kapa Biosystems). PCR was performed using an Applied Biosystems PRISMTM 7900 sequence detection system (Applied Biosystems). Specific primers were designed for detection of the *e2f1*, *e2f1-ha*, *cwp1*, *cwp2*, *cwp3*, *myb2*, thymidine kinase, *ran*, and 18S ribosomal RNA genes as follows: E2F1realF (TCCCACCCCCTCCATCA) and E2F1realR (TCCACGCATCCCTTGCA); E2F1HAF and HAR; *cwp1realF* (AACGCTCTCACAGGCTCCAT) and *cwp1realR* (AGGTG-GAGCTCCTTGAGAAATTG); *cwp2realF* (TAGGCTGCTT-CCCACCTTTGAG) and *cwp2realR* (CGGGCCCGCAA-GGT); *cwp3realF* (GCAAATTGGATGCCAAACAA) and *cwp3realR* (GACTCCGATCCAGTCGCAGTA); *myb2realF* (TCCCTAATGACGCCAAACG) and *myb2realR* (AGCAGC-CAGAGGCCAAGT); *tkrealF* (CCCCTGCGCCGTTCTC) and *tkrealR* (CCCACCGATAACAATCTCTTCTG); *ranrealF* (TCGTCCTCGTCGGAAACAA) and *ranrealR* (AACTGTCT-GGGTGCGGATCT); 18SrealF and 18SrealR. Two independently generated stably transfected lines were made from each construct, and each of these cell lines was assayed three separate times. The results are expressed as relative expression level over control. Student's *t* tests were used to determine statistical significance of differences between samples.

Plasmid Construction—All constructs were verified by DNA sequencing with a BigDye Terminator 3.1 DNA sequencing kit and a 3100 DNA Analyzer (Applied Biosystems). Plasmid 5'Δ5N-Pac was a gift from Dr. Steven Singer and Dr. Theodore Nash (60). Plasmid pNW1L has been described previously (57). For constructing pPE2F1, a PCR with oligonucleotides TU5NhF (GGCGGCTAGCCGCAGACGCATGAACGC-ATG) and TU5BaR (GGCGGGATCCTTTTATTCCGCCCGTCCAG) generated a 0.3-kb product that was digested with NheI and BamHI. Another PCR with primers E2F1BF (GGCG-GGATCCATGCAGCTCGCAGGCCAAAG) and E2F1MR (GGCGACGCGTGTGTGCTCCGTTTGCCTG) generated a 1.3-kb PCR product that was digested with BamHI and MluI and cloned into NheI/MluI-digested pPop2NHA (28) with the 0.3-kb NheI/BamHI fragment. The resulting pPE2F1 contains an *e2f1* gene controlled by the α2-tubulin promoter with an HA tag fused at its C terminus. To make the pPE2F1m1 plasmid, the *e2f1* gene was amplified using two primers E2F1m1F (GGCG-

GGATCCATGCAGCTCGCAGGCCAAAGTCCTCAAGTA-atgCAAggcAACCCGgtgTTaGttgtCTCGTCGATCTC, mutated nucleotides are shown in lowercase type) and E2F1MR. The product was digested with BamHI and MluI and cloned into the BamHI/MluI-digested pPE2F1 vector to generate plasmid pPE2F1m1. To make the pPE2F1m2 plasmid, the *e2f1* gene was amplified using two primer pairs E2F1BF and E2F1m2R (CCTATAAAAAGCacaactcatacagccAGGCGGATAGCA) and E2F1m2F (TGCTATCCGCCTggctgtatgagttgtGCTTTTTAT-AGG) and E2F1MR. The two PCR products were purified and used as templates for a second PCR. The second PCR also included primers E2F1BF and E2F1MR, and the product was digested with BamHI and MluI and cloned into the BamHI/MluI-digested pPE2F1 vector to generate plasmid pPE2F1m2. To make the pPE2F1dd plasmid, the *e2f1* gene was amplified using two primer pairs E2F1BF and E2F1dd1R (TGATGA-TAATACCGAGTATCAGATCGTAGAGC) and E2F1dd1F (GCTCTACGATCTGATACTCGGTATTATCATCA) and E2F1MR. The two PCR products were purified and used as templates for a second PCR. The second PCR also included primers E2F1BF and E2F1MR, and the product was digested with BamHI and MluI and cloned into the BamHI/MluI-digested pPE2F1 vector to generate plasmid pPE2F1dd. To make the pPE2F1DB1m (or pPE2F1DB2m) plasmid, the *e2f1* gene was amplified using two primer pairs E2F1BF and E2F1DB1mR (TATCAGATCGTAGAGCGCTCTACTTTCAATAGC) (or E2F1DB2mR, GGTAATGTCATACAGTGCACGCACCTGACTCGA) and E2F1DB1mF (GCTATTGAAAGTAGAgcGCTCTACGATCTGATA) (or E2F1DB2mF, TCGAGTCAGGTGCGTgcACTGTATGACA-TTACC) and E2F1MR. The two PCR products were purified and used as templates for a second PCR. The second PCR also included primers E2F1BF and E2F1MR, and the product was cloned into the BamHI/MluI-digested pPE2F1 vector to generate plasmid pPE2F1DB1m (or pPE2F1DB2m). The 287-bp 5'-flanking region of the thymidine kinase gene was amplified with primers TKNhe5F (GGCGGCTAGCATGGCATTTC-TATAACCTGA) and TKNco5R (GGCGCCATGGCCAATT-TTTTTTCGCGGAAAA), digested with NheI and NcoI, and ligated in place of the NheI/NcoI-excised α2-tubulin promoter sequence in pNT5 (21). The resulting plasmid, pNTK5, contained the luciferase gene under the control of the thymidine kinase promoter. The 287-bp 5'-flanking region of the thymidine kinase gene was amplified with primers TKNhe5F (GGCGGCTAGCATGGCATTTC-TATAACCTGA) and TKNco5m1R (GGCGCCATGGCCAATTTTTTTTatattgggA-TGCTG) or TKNco5m2R (GGCGCCATGGCCAATTTTTT-TTCGCGGgggATGCTGGG), digested with NheI and NcoI, and ligated in place of the NheI/NcoI-excised α2-tubulin promoter sequence in pNT5 (21). The resulting plasmid,

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

pNTK5m1 or pNTK5m2, contained the luciferase gene under the control of the thymidine kinase promoter with a mutation on the E2F1-binding site. To make the pNW1Lm1 plasmid, the 402-bp 5'-flanking region of the *cwp1* gene was amplified using two primer pairs *cwp15NF* and *w1m1R* (GACTGTAGACTGTTAatactcccAAATGCAGAC) and *w1m1F* (GTCTGCATTTgggagttTAACAGTCTACAGTC) and *cwp15NR* (GGCGGC-CATGGCCCTGATATTTTATTCTGTGTT). The two PCR products were purified and used as templates for a second PCR. The second PCR also included primers *cwp15NF* and *cwp15NR*, and the product was digested with *NheI* and *NcoI* and ligated in place of the *NheI/NcoI*-excised α -tubulin promoter sequence in pNT5 (21). The resulting plasmid, pNW1Lm1, contained the luciferase gene under the control of the *cwp1* promoter with a mutation on the putative E2F1-binding site. The 402-bp 5'-flanking region of the *cwp1* gene was amplified from pPW1m4 (24) template using two primers *cwp15NF* (GgCGgtagcttcgcttgctttgCGgtagaagag) and *w1m2NR* (GGCGCCATGGCCCTGGcgcggcccccTGTGTTCTTGATCTGAGAG), digested with *NheI* and *NcoI*, and ligated in place of the *NheI/NcoI*-excised α -tubulin promoter sequence in pNT5 (21). The resulting plasmid, pNW1Lm2, contained the luciferase gene under the control of the *cwp1* promoter with a mutation on the AT-rich Inr element.

5'-RACE Analysis—5'-RACE for determination of the transcription start sites of the thymidine kinase gene was performed using the 5'-RACE system (Invitrogen). Oligonucleotides TKGSPR1 (TCGCCAAACTCTGTTGCAAGT) and TKGSPR2 (TCGCCAAGAAGTACCGCTCTC) were used as first-strand primer and nested primer respectively.

Transfection, Luciferase Assay, and Western Blot Analysis—Cells transfected with the pP series plasmid containing the *pac* gene were selected and maintained with 54 μ g/ml puromycin. The luciferase activity was determined as described previously (17). After stable transfection with specific constructs, luciferase activity was determined in vegetative cells at late log/stationary phase (1.5×10^6 cells/ml) as described previously (17) and was measured with an Optocomp I luminometer (MGM Instruments). Two independently generated stably transfected lines were made from each construct, and each of these lines was assayed three separate times. Western blots were probed with anti-V5 horseradish peroxidase (HRP) (Invitrogen) or anti-HA monoclonal antibody (Covance, Princeton, NJ; 1:5000 in blocking buffer), anti-CWP1 (1:10,000 in blocking buffer) (26), anti-E2F1 (1:10,000 in blocking buffer) (see below), or pre-immune serum (1:5000 in blocking buffer), and detected with peroxidase-conjugated goat anti-mouse IgG (Pierce, 1:5000) or peroxidase-conjugated goat anti-rabbit IgG (Pierce, 1:5000) and enhanced chemiluminescence (GE Healthcare).

Expression and Purification of Recombinant E2F1 Protein—The genomic *e2f1* gene was amplified using oligonucleotides E2F1F and E2F1R. The product was cloned into the expression vector pET101/D-TOPO (Invitrogen) in-frame with the C-terminal His and V5 tag to generate plasmid pE2F1. To make the pE2F1m1, pE2F1m2, pE2F1dd, pE2F1DB1m, or pE2F1DB2m expression vector, the *e2f1* gene was amplified using primers E2F1F and E2F1R and specific template, including pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, or pPE2F1DB2m. The

product was cloned into the expression vector to generate plasmid pE2F1m1, pE2F1m2, pE2F1dd, pE2F1DB1m, or pE2F1DB2m. The pE2F1, pE2F1m1, pE2F1m2, pE2F1dd, pE2F1DB1m, or pE2F1DB2m plasmid was freshly transformed into *Escherichia coli* BL21 StarTM(DE3) (Invitrogen). An overnight pre-culture was used to start a 250-ml culture. *E. coli* cells were grown to an A_{600} of 0.5 and then induced with 1 mM isopropyl thiogalactopyranoside (Promega) for 4 h. Bacteria were harvested by centrifugation and sonicated in 10 ml of buffer A (50 mM sodium phosphate, pH 8.0, 300 mM NaCl) containing 10 mM imidazole and protease inhibitor mixture (Sigma). The samples were centrifuged, and the supernatant was mixed with 1 ml of 50% slurry of nickel-nitrilotriacetic acid Superflow (Qiagen). The resin was washed with buffer A containing 20 mM imidazole and eluted with buffer A containing 250 mM imidazole. Fractions containing E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m were pooled, dialyzed in 25 mM HEPES, pH 7.9, 20 mM KCl, and 15% glycerol, and stored at -70°C . Protein purity and concentration were estimated by Coomassie Blue and silver staining compared with serum albumin. E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m was purified to apparent homogeneity (>95%).

Generation of Anti-E2F1 Antibody—Purified E2F1 protein was used to generate rabbit polyclonal antibodies through a commercial vendor (Angene, Taipei, Taiwan).

Generation of Anti-Myb2 Antibody—The genomic *myb2* gene was amplified using oligonucleotides Myb2F and Myb2R, and the product was cloned into the expression vector pCART7/CT-TOPO (Invitrogen) in-frame with the C-terminal His and V5 tags to generate plasmid pMyb. The pMyb plasmid was freshly transformed into *E. coli* BL21(DE3)pLysS (QIAexpressionist, Qiagen). An overnight pre-culture was used to start a 250-ml culture. *E. coli* cells were grown to an A_{600} of 0.5, and then induced with 1 mM isopropyl thiogalactopyranoside (Promega) for 2 h. Bacteria were harvested by centrifugation and sonicated in 10 ml of buffer A (100 mM sodium phosphate, 10 mM Tris-Cl, 6 M guanidine hydrochloride, pH 8.0) containing 10 mM imidazole and complete protease inhibitor mixture (Roche Applied Science). The samples were centrifuged, and the supernatant was mixed with 1 ml of a 50% slurry of nickel-nitrilotriacetic acid superflow (Qiagen). The resin was washed with buffer B (100 mM sodium phosphate, 10 mM Tris-Cl, 8 M urea, pH 8.0) and buffer C (100 mM sodium phosphate, 10 mM Tris-Cl, 8 M urea, pH 6.3) and eluted with buffer E (100 mM sodium phosphate, 10 mM Tris-Cl, 8 M urea, pH 4.5). Fractions containing Myb2 were pooled, dialyzed in 25 mM HEPES, pH 7.9, 40 mM KCl, 0.1 mM EDTA, and 15% glycerol, and stored at -70°C . Protein purity and concentration were estimated by Coomassie Blue and silver staining compared with bovine serum albumin. Myb2 was purified to apparent homogeneity (>95%). Purified Myb2 protein was used to generate rabbit polyclonal antibodies through a commercial vendor (Angene, Taipei, Taiwan).

GST Pulldown Assay—To clone pGSTMyb2, a PCR product generated by oligonucleotides GMEF (GGCGGAATTCATGTTACCGGTACCTTCTCA) and GMSR (GGCGGTCGACTCAGGGTAGCTTCTCACGGG) on genomic DNA template was digested with *EcoRI* and *Sall* and cloned into *EcoRI/Sall*-

digested pGEX-6P-1 (Amersham Biosciences). The pGEX6p-1 or pGSTMyb2 plasmid was transformed into *E. coli* BL21(DE3)pLysS. Expression was induced by addition of isopropyl thiogalactopyranoside to a final concentration of 1 mM and continued culture growth for 2 h. GST or GSTMyb2 fusion proteins expressed in *E. coli* were affinity-purified with the Microspin GST purification module according to the manufacturer's instructions (GE Healthcare).

To confirm the presence of GST and GSTMyb2 in the beads, the beads were subjected to SDS-PAGE. GST and GSTMyb2 were visualized by Coomassie Blue staining. For GST pull-down assay, 5–10 μ g of GST or GSTMyb2 was incubated with 50 ng of purified recombinant E2F1 protein with a C-terminal V5 tag in a pull-down buffer (25 mM Tris phosphate, pH 7.8, 2 mM DTT, 10% glycerol, 1% Triton X-100, 1% Nonidet P-40, 300 mM KCl). After incubation for 1 h at room temperature, the beads were washed five times in the pull-down buffer. Finally the beads were resuspended in SDS-PAGE sample buffer and analyzed by Western blot and probed with anti-V5-HRP antibody (Invitrogen).

Co-immunoprecipitation Assay—The 5'Δ5N-Pac and pPE2F1 stable transfectants were inoculated into encystation medium (5×10^7 cells in 45 ml of medium) and harvested after 24 h in encystation medium under drug selection and washed in phosphate-buffered saline. Cells were lysed in luciferase lysis buffer (Promega) and protease inhibitor (Sigma) and then vortexed with glass beads. The cell lysates were collected by centrifugation and then incubated with anti-HA antibody conjugated to beads (Bethyl Laboratories Inc.). The beads were washed three times with luciferase lysis buffer (Promega). Finally, the beads were then resuspended in sample buffer and analyzed by Western blot and probed with anti-HA monoclonal antibody (1:5000 in blocking buffer; Sigma) and anti-Myb2 (1:10,000 in blocking buffer) and detected with peroxidase-conjugated goat anti-mouse IgG (Pierce, 1:5000) or peroxidase-conjugated goat anti-rabbit IgG (Pierce, 1:5000) and enhanced chemiluminescence (GE Healthcare).

Immunofluorescence Assay—The pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, or pPE2F1DB2m stable transfectants were cultured in growth medium under puromycin selection. Cells cultured in growth medium or encystation medium for 24 h were harvested, washed in phosphate-buffered saline (PBS), attached to glass coverslips (2×10^6 cells/coverslip), and then fixed and stained (17). Cells were reacted with anti-HA monoclonal antibody (1:300 in blocking buffer; Molecular Probes) and anti-mouse Alexa 488 (1:500 in blocking buffer, Molecular Probes) as the detector. For staining cysts, cysts were reacted with anti-CWP1 (1:300 in blocking buffer) and anti-rabbit Alexa 488 (Molecular Probes, 1:500 in blocking buffer) was used as the detector. ProLong antifade kit with 4',6-diamidino-2-phenylindole (Invitrogen) was used for mounting. E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m was visualized using a Leica TCS SP5 spectral confocal system.

Electrophoretic Mobility Shift Assay—Double-stranded oligonucleotides specified throughout were 5'-end-labeled as described previously (18). Binding reaction mixtures contained the components described previously (24). Labeled probe (0.02 pmol) was incubated for 15 min at room temperature with 5

ng of purified E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m protein in a 20- μ l volume supplemented with 0.5 μ g of poly(dI-dC) (Sigma). Competition reactions contained 200-fold molar excess of cold oligonucleotides. In an antibody supershift assay, 0.8 μ g of an anti-V5-HRP antibody (Bethyl Laboratories) was added to the binding reaction mixture. The mixture was separated on a 6% acrylamide gel by electrophoresis.

ChIP Assays—The WB clone C6 cells were inoculated into encystation medium (5×10^7 cells in 45 ml medium) and harvested after 24 h in encystation medium under drug selection and washed in phosphate-buffered saline. ChIP was performed as described previously (26) with some modifications. Formaldehyde was then added to the cells in phosphate-buffered saline at a final concentration of 1%. Cells were incubated at room temperature for 15 min, and reactions were stopped by incubation in 125 mM glycine for 5 min. After phosphate-buffered saline washes, cells were lysed in luciferase lysis buffer (Promega) and protease inhibitor (Sigma) and then vortexed with glass beads. The cell lysate was sonicated on ice and then centrifuged. Chromatin extract was incubated with protein G plus/protein A-agarose (Merck) for 1 h. After removal of protein G plus/protein A-agarose, the precleared lysates were incubated with 2 μ g of anti-E2F1 antibody or preimmune serum for 2 h and then incubated with protein G plus/protein A-agarose (Merck) for 1 h. The beads were washed twice with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.0, 150 mM NaCl), once with high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.0, 500 mM NaCl), once with LiCl buffer (0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0), and twice with TE buffer (20 mM Tris-HCl, 1 mM EDTA, pH 8.0). The beads were resuspended in elution buffer containing 50 mM Tris-HCl, pH 8.0, 1% SDS, and 10 mM EDTA at 65 °C for 4 h. To prepare DNA representing input DNA, 2.5% of precleared chromatin extract without incubation with anti-E2F1 was combined with elution buffer. Eluted DNA was purified by the QIAquick® PCR purification kit (Qiagen). Purified DNA was subjected to PCR followed by agarose gel electrophoresis. Primers 18S5F (CCAAAAAAGTGTGGTGCAGG) and 18S5R (GCCGGGCGCGGGCGCCGCGG) were used to amplify the 18S ribosomal RNA gene promoter as a control for our ChIP analysis. Primers E2F15F (GAAAGTTTAATCAGGCACAT) and E2F15R (TAAAAAATGAATTGCGGGAA), *cwp15F* (CAACGGCTTACTAAATCATTCTCTTG) and *cwp15R* (TTCTGTGTTTCTTGATCTGAGAGTTGT), *cwp25F* (GAGCATGGTTCTGACAAATAGTG) and *cwp25R* (CATCAGTCTACTGTTTCTTTTGTAGTTTCATATCT), *cwp35F* (TTCGGGAGGGTCTGGGGGAG) and *cwp35R* (ATCAGTAGTAACTT-ATTTTT), *myb25F* (TGCACTGTAGCGTTTCCATTTG) and *myb25R* (ACTTACCCGTAATGGCGTTGAC), *tk5F* (AGCAAGACATTAGTAATGCT) and *tk5R* (CCAATTTTT-TTTCGCGGAAAA), and *ran5F* (GCCGCTTCAATGACAG-ATGC) and *ran5R* (GCTACTCTCGGTTCCCTGGGT) were used to amplify *e2f1*, *cwp1*, *cwp2*, *cwp3*, *myb2*, thymidine kinase, and *ran* gene promoters within the –200 to –1 region.

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

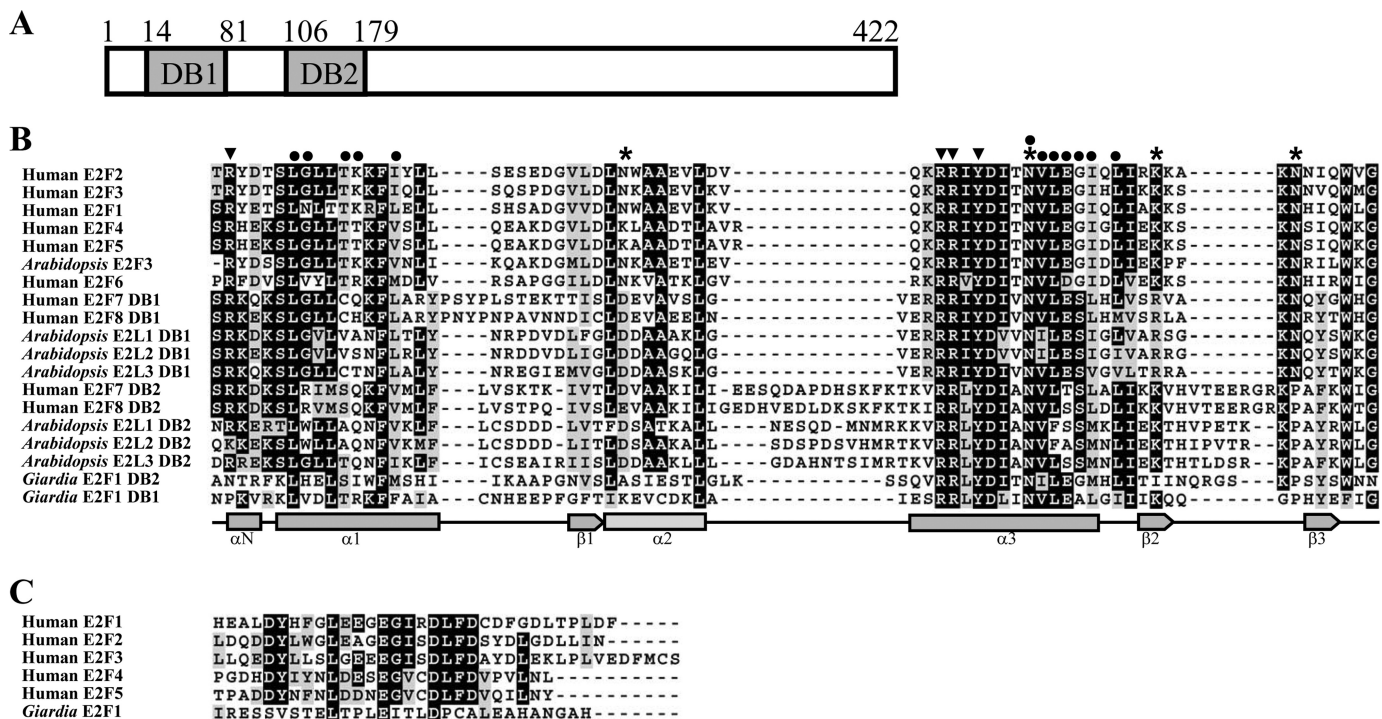


FIGURE 1. Domain architecture of E2F1 protein and alignment of the DNA-binding domains of E2Fs. *A*, schematic representation of the giardial E2F1 protein. The gray boxes indicate the two DNA-binding domains (DB1 and DB2). *B*, alignment of the DNA-binding domains of E2Fs. The DNA-binding domains from members of the E2F family are analyzed by ClustalW 1.83 (86), including human E2F1, E2F2, E2F3, E2F4, E2F5, E2F6, E2F7, and E2F8 (accession numbers are AAC50719, NP_004082, NP_001940, NP_001941, Q15329, O75461, NP_976328, and EAW68354, respectively); *Arabidopsis* E2F3, E2L1, E2L2, and E2L3 (accession numbers are NP_973611, BAB91412, BAB91413, and BAB91414, respectively); and putative E2F1 from *G. lamblia* (GenBank™ accession numbers XP_001705587, open reading frames 23756 in the *G. lamblia* genome database). These E2Fs contain either one or two DNA-binding domains (DB1 and DB2). Letters in black boxes, letters in gray boxes, and hyphens indicate identical amino acids, similar amino acids, and gaps in the respective proteins, respectively. Residues in human E2F4 (49) making heterodimerization contact, DNA base contact, and DNA backbone contact are indicated by circles, arrowheads, and asterisks, respectively. *C*, alignment of the pRB-binding domains of E2Fs. The C-terminal pRB-binding domains of the human E2F proteins and the C-terminal end of putative E2F1 from *G. lamblia* are analyzed by ClustalW 1.83 (67, 86).

Microarray Analysis—RNA was quantified by A_{260} nm by an ND-1000 spectrophotometer (Nanodrop Technology) and qualified by a Bioanalyzer 2100 (Agilent Technology) with an RNA 6000 Nano LabChip kit. RNA from the pE2F1 cell line was labeled by Cy5 and RNA from the 5' Δ 5N-Pac cell line was labeled by Cy3. 0.5 μ g of total RNA was amplified by a Low RNA Input Fluor Linear Amp kit (Agilent Technologies) and labeled with Cy3 or Cy5 (CyDye, PerkinElmer Life Sciences) during the *in vitro* transcription process. 0.825 μ g of Cy-labeled cRNA was fragmented to an average size of about 50–100 nucleotides by incubation with fragmentation buffer at 60 °C for 30 min. Correspondingly fragmented labeled cRNA was then pooled and hybridized to a *G. lamblia* oligonucleotide microarray (Agilent Technologies) at 60 °C for 17 h. After washing and drying by nitrogen gun blowing, microarrays were scanned with an Agilent microarray scanner (Agilent Technologies,) at 535 nm for Cy3 and 625 nm for Cy5. Scanned images were analyzed by Feature Extraction version 9.1 software (Agilent Technologies), and image analysis and normalization software was used to quantify signal and background intensity for each feature; data were substantially normalized by the rank consistency filtering LOWESS method.

RESULTS

Identification and Characterization of *e2f1* Gene—To identify genes encoding novel E2F1 proteins from *G. lamblia*, we

performed BLAST searches (59) against the *G. lamblia* genome database (11, 58) using the amino acid sequences of the human E2F1 (GenBank™ accession numbers U47675 and U47646) as a query sequence. This search detected one putative homolog, E2F1 (GenBank™ accession number XM_001705535, open reading frame 23756 in the *G. lamblia* genome database). Comparison of genomic and cDNA sequences showed that the *e2f1* gene contained no introns. The deduced giardial E2F1 protein contains 422 amino acids with a predicted molecular mass of ~46.22 kDa and a pI of 8.22. It has two putative E2F/DP family winged-helix DNA-binding domains as predicted by Pfam (61). Like the human E2F family (30), the DNA-binding domains of which are near the N-terminal region, the two DNA-binding domains of the giardial E2F1 are near the N-terminal region (residues 14–81 and 106–179) (Fig. 1A). The similarity between E2F1 and other E2F proteins is limited to their DNA-binding domains (data not shown).

The structure of the DNA-binding domain of E2F includes four α -helices and three β -sheets, which form a hydrophobic core (Fig. 1B) (49). Structural studies of the DNA-binding domain of human E2F4 show that the α 3 helix binds in the major groove and the two Arg of RRYD (arrowheads in Fig. 1B) are the key residues contacting the base G of the binding sequence (TTTCGCGCG) on opposite strands (49). The Tyr residue of RRYD (arrowhead in Fig. 1B) contacts a base C of

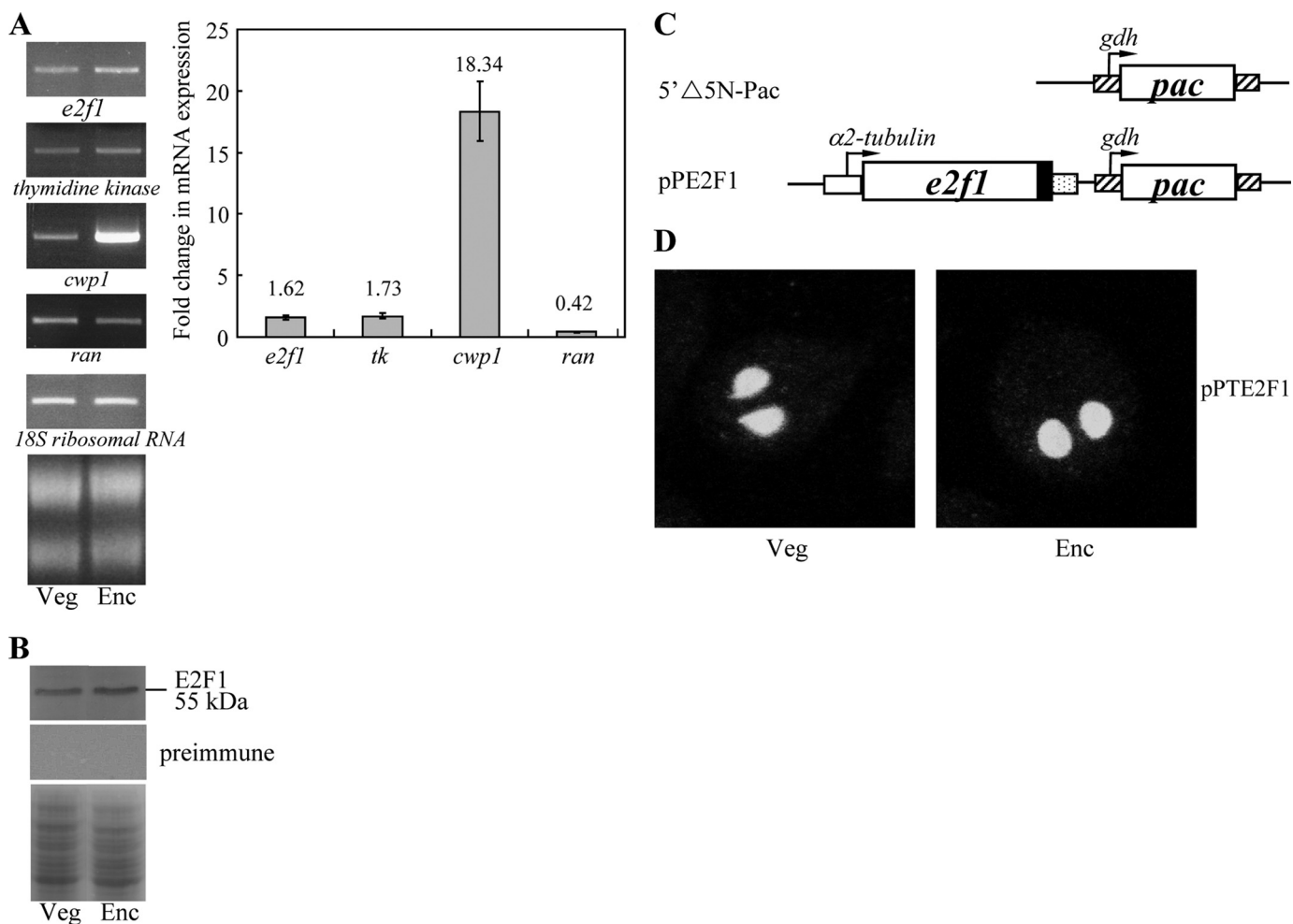


FIGURE 2. Analysis of *e2f1* gene expression. A, RT-PCR and quantitative real time PCR analysis of *e2f1* gene expression. RNA samples were prepared from *G. lamblia* wild type nontransfected WB cells cultured in growth (Veg, vegetative growth) or encystation medium and harvested at 24 h (Enc, encystation). RT-PCR was performed using primers specific for *e2f1*, thymidine kinase, *cwp1*, *ran*, and 18 S ribosomal RNA genes. Ribosomal RNA quality and loading controls are shown in the bottom panel. Representative results are shown on left. Real time PCR was performed using primers specific for *e2f1*, thymidine kinase, *cwp1*, *ran*, and 18 S ribosomal RNA genes. Transcript levels were normalized to 18 S ribosomal RNA levels. Fold changes in mRNA expression are shown as the ratio of transcript levels in encysting cells relative to vegetative cells. Results are expressed as the means \pm S.E. of at least three separate experiments (right panel). B, E2F1 protein levels in different stages. The wild type nontransfected WB cells were cultured in growth (Veg, vegetative growth) or encystation medium for 24 h (Enc, encystation) and then subjected to SDS-PAGE and Western blot. The blot was probed by anti-E2F1 antibody or preimmune serum. Representative results are shown. Equal amounts of protein loading were confirmed by SDS-PAGE and Coomassie Blue staining. C, diagrams of the 5'Δ5N-Pac and pPE2F1 plasmid. The *pac* gene (open box) is under the control of the 5' - and 3' -flanking regions of the *gdh* gene (striated box). In construct pPE2F1, the *e2f1* gene is under the control of the 5' -flanking region of the constitutively expressed $\alpha 2$ -tubulin promoter (open box) and the 3' -flanking region of the *ran* gene (dotted box). The filled black box indicates the coding sequence of the HA epitope tag. D, nuclear localization of E2F1. The pPE2F1 stable transfectants were cultured in growth (Veg, vegetative growth, left panels) or encystation medium for 24 h (Enc, encystation, right panels) and then subjected to immunofluorescence analysis using anti-HA antibody for detection. The product of pPE2F1 localizes to the nuclei in both vegetative and encysting trophozoites.

the binding sequence (49). The Asp residue of the RRXYD motif helps stabilize the two Args by making hydrogen bonds to them (49). Several amino acids of the α -helices and β -sheets contact the DNA backbone (asterisks in Fig. 1B) (49). The Arg of the α N helix (arrowhead in Fig. 1B) contacts the DNA minor groove near the T-rich region of the binding sequence. The $\alpha 1$ and $\alpha 3$ helices make dimerization contact by forming a hydrophobic interface (circles in Fig. 1B). Sequence alignment of the two DNA-binding domains of the giardial E2F1 shows that some residues making DNA base contact, including the RRXYD motifs of the $\alpha 3$ helix, are conserved; the Arg of the α N helix making DNA base contact is not conserved (arrowhead in Fig. 1B); some residues making DNA backbone contact or dimerization contact are conserved (asterisks or circles in Fig. 1B). The full-length of giardial E2F1 has 13% identity and 28%

similarity to that of human E2F4, but it has lower identity (10%) and similarity (24%) to that of human DP1. The DNA-binding domain 1 (2) of giardial E2F1 has 34% (30%) identity and 46% (54%) similarity to the DNA-binding domain of the human E2F4. The C-terminal 69 amino acids of the human E2F1 have been reported to be a transactivation domain and pRB-binding domain (42, 62). The C-terminal half of the giardial E2F1 (residues 180–422), which does not contain the DNA-binding domains, has no apparent functional motif. The C-terminal end of the giardial E2F1 is not highly similar to the pRB-binding domain of the human E2Fs (Fig. 1C).

Expression of the *e2f1* Gene—RT-PCR and quantitative real time PCR analysis of total RNA showed that the *e2f1* transcript was present in vegetative cells and increased significantly in 24-h encysting cells (Fig. 2A). As controls, we found that the

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

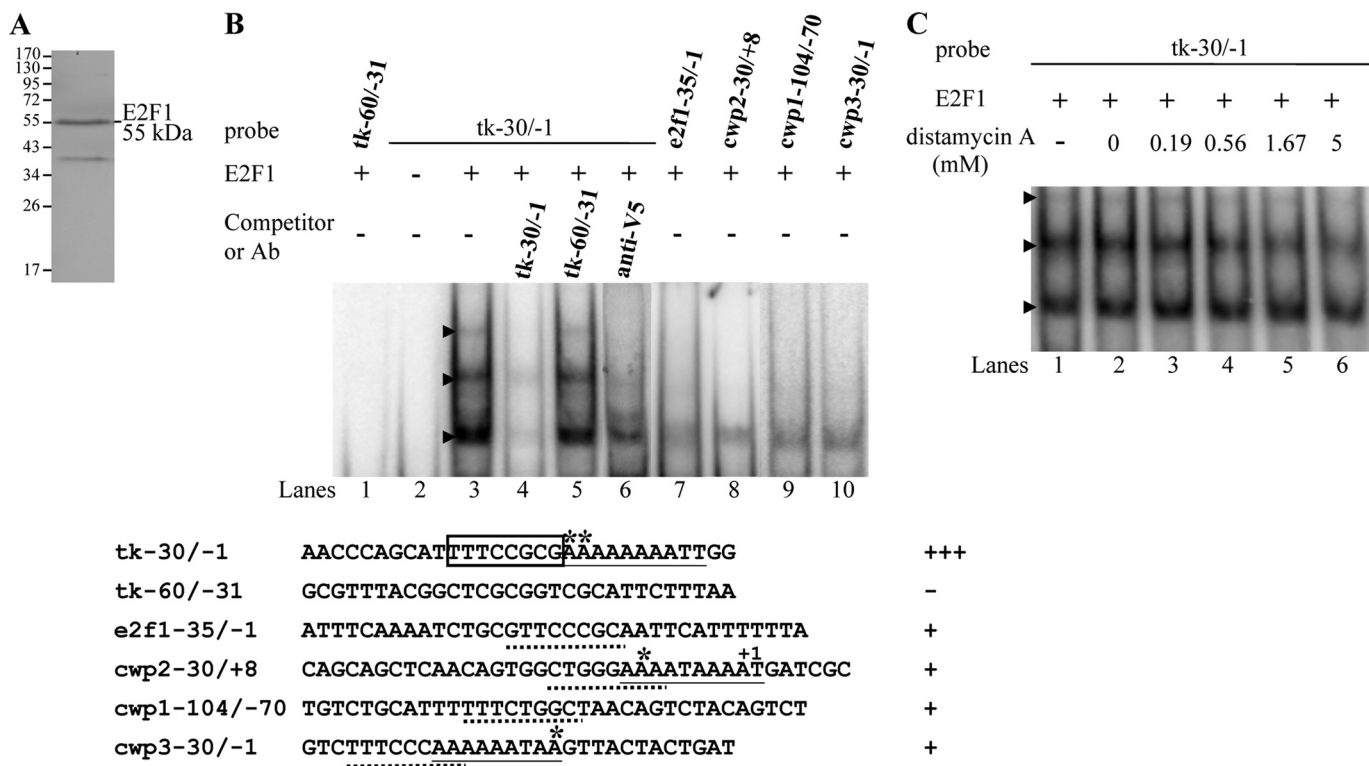


FIGURE 3. DNA binding ability of E2F1 revealed by electrophoretic mobility shift assays. *A*, Western blot analysis of recombinant E2F1 protein with a V5 tag at its C terminus purified by affinity chromatography. The purified E2F1 protein is detected by anti-V5-HRP antibody. The lower molecular weight band was a degraded form. *B*, detection of E2F1-binding sites. Electrophoretic mobility shift assays were performed using purified E2F1 and ³²P-end-labeled oligonucleotide probes tk-30/-1, tk-60/-31, or other probes as described. Numbers of the 5'-flanking region of the thymidine kinase (*tk*) and *cwp* genes are relative to the translation start site (+1). Components in the binding reaction mixtures are indicated *above* the lanes. The E2F1 binding specificity for tk-30/-1 probe was confirmed by competition and supershift assays. Some reaction mixtures contained 200-fold molar excess of cold oligonucleotides tk-30/-1 or tk-60/-1 or 0.8 μg of anti-V5-HRP antibody, as indicated *above* the lanes. The transcription start sites of the thymidine kinase gene determined from vegetative cells are indicated by *asterisks*. The transcription start sites of the *cwp2* and *cwp3* gene determined from 24-h encysting cells (15, 23) are indicated by *asterisks*. The AT-rich Inr element spanning the transcription start sites is *underlined*. The putative E2F1-binding sequence of the thymidine kinase gene is *framed*. The putative E2F1-binding sequence of the *e2f1*, *cwp1*, *cwp2*, or *cwp3* promoter is *underlined*. + + +, +, and - represent strong binding, weak binding, and no binding, respectively. *C*, effect of distamycin A on the binding of E2F1 to DNA. ³²P-End-labeled tk-30/-1 probe was incubated with E2F1 in the absence (*lane 1*) or presence of distamycin A (*lanes 3-6*). The *arrowheads* indicate the shifted complexes. Distamycin A was dissolved in Me₂SO. Adding Me₂SO to the reaction mixture did not decrease the E2F1 binding activity (*lane 2*).

mRNA levels of the *cwp1* and *ran* genes increased and decreased significantly during encystation, respectively (Fig. 2A). The products of the *cwp1* and *ran* genes are the component of the cyst wall and the *ras*-related nuclear protein (15, 18). To determine the expression of E2F1 protein, we generated an antibody specific to the full-length E2F1. Western blot analysis confirmed that this antibody recognized E2F1 at a size of ~55 kDa (Fig. 2B), which was almost matched to the predicted molecular mass of E2F1 (~46.2 kDa). E2F1 was expressed in vegetative cells, and its levels increased significantly during encystation (Fig. 2B). The preimmune serum did not detect any bands at a size of ~55 kDa (Fig. 2B).

Localization of the E2F1 Protein—To determine the role of E2F1 protein in *G. lamblia*, we expressed *e2f1* gene constitutively under the control of the α2-tubulin gene promoter with an HA epitope tag at its C terminus (Fig. 2C) and observed its expression in *Giardia*. The HA-tagged E2F1 was detected in the nuclei during vegetative growth and encystation (Fig. 2D and [supplemental Fig. S1](#)), indicating that E2F1 is a nuclear protein in *Giardia*. We also tried to express the *ef21* gene under its own promoter with an HA epitope tag at its C terminus (data not shown). The expression of the HA-tagged E2F1 protein can be detected by Western blot analysis and immunofluorescence

assays, but the signal was too weak to be shown (data not shown). E2F1 protein was also detected in the nuclei using an anti-E2F1 antibody for immunofluorescence assays (data not shown).

Binding of E2F1 to Thymidine Kinase Promoter—The nuclear localization of E2F1 suggested that it might also function as a transcription factor in *G. lamblia*. To test its DNA binding activity, we expressed E2F1 with a C-terminal V5 tag in *E. coli* and purified it to >95% homogeneity (data not shown). An anti-V5-HRP antibody specifically recognized the recombinant V5-tagged E2F1 (~55 kDa) in Western blot (Fig. 3A). A minor band (~40 kDa) that may be a degraded form was also detected (Fig. 3A).

Electrophoretic mobility shift assays were performed with the purified E2F1 and double-stranded DNA sequences from the 5'-flanking region of thymidine kinase gene, an S phase-specific gene (63). Incubation of a labeled double-stranded DNA probe tk-30/-1 with E2F1 resulted in the formation of retarded bands (Fig. 3B, *lane 3*). tk-30/-1 is the region from -30 to -1 bp relative to the translation start site of the thymidine kinase gene. E2F1 did not bind to either single strand of the tk-30/-1 probe (data not shown). Incubation of the labeled probe tk-60/-31 with E2F1 did not form any retarded bands

(Fig. 3B, lane 1). The binding specificity was confirmed by competition and supershift assays (Fig. 3B, lanes 4–6). The formation of the shifted tk–30/–1 bands was almost totally competed by a 200-fold molar excess of unlabeled tk–30/–1 but not by the same excess of a nonspecific competitor, tk–60/–31 (Fig. 3B, lanes 4 and 5). Addition of the anti-V5-HRP antibody that recognized the purified E2F1 decreased the levels of the normal mobility shift complex (Fig. 3B, lane 6). The results suggest that *Giardia* E2F1 can bind the thymidine kinase promoter (–30/–1 region) and that it can bind DNA independently of DP, like other E2Fs with two DNA-binding domains (43, 53, 55). Interestingly, the tk–30/–1 probe contains a TTTCCGCG sequence that is similar to the consensus human E2F-binding sites TTT(c/g)GCGC(c/g) or TTTCCCGCC (42, 45–52).

E2F1 was also shown to bind to its own promoter (e2f1–35/–1 probe) (Fig. 3B, lane 7). We also tested whether E2F1 binds to the promoter of encystation-induced genes. We found that E2F1 also bound to encystation-specific *cwp1*, *cwp2*, and *cwp3* promoter (*cwp1*–104/–70, *cwp2*–30/+8, and *cwp3*–30/–1 probes) (Fig. 3B, lanes 8–10). The results suggest that these specific promoters may contain putative E2F1-binding sequences (Fig. 3B, underlined letters). E2F1 did not bind to the –30 to –1 and –60 to –31 region of the 5′-flanking region of the 18 S ribosomal RNA gene, which do not contain the putative E2F1-binding sequences (data not shown).

Studies suggest that human E2F4 can bind to both DNA major and minor grooves (arrowheads in Fig. 1B) (49). To investigate how E2F1 binds DNA, we used distamycin A, which binds to the DNA minor groove, as a competitive inhibitor of E2F1 binding (64). As shown in Fig. 3C, the binding of E2F1 to DNA decreased with increasing concentrations of distamycin A. However, the binding was not completely inhibited at concentrations ~5 mM, suggesting that E2F1 may bind to both major and minor grooves.

Scanning mutagenesis of the tk–30/–1 probe showed that substitutions within the TTTCCGCG sequence significantly decreased the DNA-protein interaction (tk–30/–1m1–3 and tk–30/–1–1m5; Fig. 4A, lanes 2–4 and 6), but mutations of the other regions caused a minor decrease in binding (tk–30/–1–1m7–9 and tk–30/–1–1m11–12; Fig. 4A, lanes 9–11, 13, and 14). Mutation of the AT-rich Inr sequence located downstream of the TTTCCGCG sequence also caused a decrease in binding (tk–30/–1–1m4, –6 and –10; Fig. 4A, lanes 5, 7, and 12). Scanning mutagenesis of the TTTCCGCG sequence in the tk–30/–1 probe showed that any single substitution within the sequence decreased the binding significantly (tk–30/–1m13–17 and tk–30/–1m20–23; Fig. 4A, lanes 16–20 and 24–27) and mutation of the fifth G and sixth C of the TTTCCGCG sequence completely eliminated the binding (tk–30/–1–30/–1m18 and tk–30/–1m19; Fig. 4A, lanes 21 and 23). The results suggest that the giardial E2F1-binding site may be TTTCCGCG, which is similar to the binding site of the known E2F1 family proteins (TTT(c/g)GCGC(c/g) or TTTCCCGCC (42, 45–52).

Recruitment of E2F1 to the *e2f1*, *cwp1*–3, and *myb2* Promoters—We further used CHIP assays to study the association of E2F1 with specific promoters in the E2F1-overexpressing cell line.

We found that E2F1 was associated with its own promoter and the *cwp1*, *cwp2*, *cwp3*, *myb2*, thymidine kinase, and *ran* promoters during vegetative growth or during encystation (Fig. 4C and data not shown). However, E2F1 was not associated with the 18 S ribosomal RNA gene promoter that has no E2F1-binding site in the <200-bp 5′-flanking region (Fig. 4C).

Transactivation of Thymidine Kinase Promoter Activity by E2F1—Identification of the E2F1-binding sequence in the thymidine kinase promoter allowed us to examine whether the E2F1 protein actually activates transcription of this promoter. Transactivation was examined by stably transfecting a reporter plasmid containing a luciferase gene into *Giardia* together with an effector plasmid pPE2F1 in which expression of the full-length E2F1 was driven by the α 2-tubulin promoter. Co-transfection of the pNTK5 reporter construct that expressed the luciferase gene under the control of the thymidine kinase promoter (Fig. 5A) with pPE2F1 resulted in an ~20-fold increase (Fig. 5B) in luciferase activity during normal growth compared with the activity in the pNTK5 + 5′ Δ 5N-Pac co-transfectants (5′ Δ 5N-Pac is the construct expressing only the puromycin selection marker) (60). To understand whether the increase in luciferase activity was due to increased mRNA levels, we measured the luciferase mRNA levels from different transfectants. Co-transfection of pPE2F1 into the pNTK5 transfectants resulted in an ~18-fold increase in the levels of the luciferase mRNA compared with the levels in the pNTK5 + 5′ Δ 5N-Pac co-transfectants (Fig. 5B), indicating that E2F1 can transactivate the thymidine kinase promoter.

We further asked whether E2F1 transactivates the thymidine kinase promoter through its binding site. The E2F1-binding sequence of the thymidine kinase promoter was mutated to construct the plasmid pNTK5m1 or pNTK5m2, and the effect of E2F1 was tested by co-transfection. We found that mutation of the E2F1-binding sequence of the thymidine kinase promoter resulted in a significant decrease of the promoter activity to ~60–80% of the wild type activity during vegetative growth (Fig. 5A). In addition, the transactivation levels in the pNTK5m1 + pPE2F1 or pNTK5m2 + pPE2F1 co-transfectants decreased significantly (from ~20- to ~5–10-fold) compared with the pNTK5 + 5′ Δ 5N-Pac cell line (Fig. 5, B–D). Therefore, mutation of the E2F1 target sequence in the thymidine kinase promoter decreased the transactivation effect by E2F1. The results suggest that E2F1 can transactivate the thymidine kinase promoter *in vivo* through its binding sequence.

Transactivation of *cwp1* Promoter Activity by E2F1—We also tried to understand the effect of E2F1 on the *cwp1* promoter that is required for encystation. E2F1 can bind to the *cwp1* promoter *in vitro* (Fig. 3B). The effect of E2F1 on the *cwp1* promoter was examined by stably transfecting a reporter plasmid containing a luciferase gene into *Giardia* together with an effector plasmid pPE2F1 (Fig. 2C). Co-transfection of the pNW1L reporter construct that expressed the luciferase gene under the control of the *cwp1* promoter (Fig. 6A) with pPE2F1 resulted in an ~5-fold increase (Fig. 6B) in luciferase activity during normal growth compared with the activity in the pNW1L + 5′ Δ 5N-Pac co-transfectants. To understand whether the increase in luciferase activity was due to increased mRNA levels, we measured the luciferase mRNA levels from

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

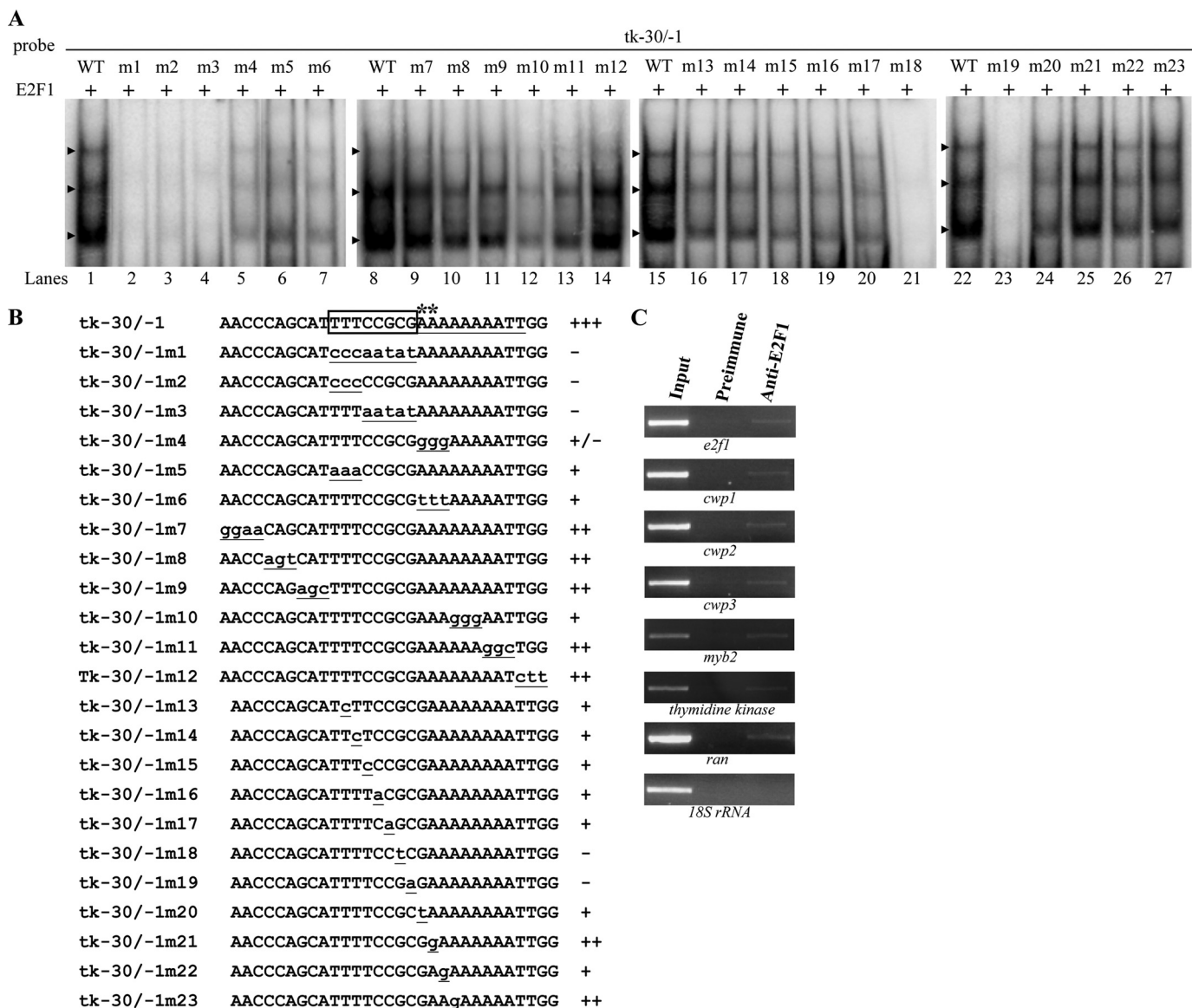


FIGURE 4. Mutation analysis of the tk-30/-1 probe sequence containing the putative E2F1-binding site. *A* and *B*, electrophoretic mobility shift assays were performed using purified E2F1 and various 32 P-end-labeled tk-30/-1 mutant probes as described in *B*. Base changes in the mutants are shown in *underlined lowercase type*. Components in the binding reaction mixtures are indicated *above* the lanes. The *arrowheads* indicate the shifted complexes. The transcription start sites of the thymidine kinase gene determined from vegetative cells are indicated by *asterisks*. The AT-rich Inr element spanning the transcription start sites is *underlined*. +, +/-, and - represent moderate binding, weak binding, and no binding, respectively. +++ and ++ represent strong binding. *C*, recruitment of E2F1 to the *cwp1-3*, *myb2*, and thymidine kinase promoters. The nontransfected WB cells were cultured in growth medium for 24 h and then subjected to ChIP assays. Anti-E2F1 was used to assess binding of E2F1 to endogenous gene promoters. Preimmune serum was used as a negative control. Immunoprecipitated chromatin was analyzed by PCR using primers that amplify the 5'-flanking region of specific genes. At least three independent experiments were performed. Representative results are shown. Immunoprecipitated products of E2F1 yielded more PCR products of *e2f1*, *cwp1-3*, *myb2*, thymidine kinase, and *ran* promoters, indicating that E2F1 was bound to these promoters. The 18 S ribosomal RNA gene promoter was used as a negative control for our ChIP analysis.

different transfectants. Co-transfection of pPE2F1 into the pNW1L transfectants resulted in an ~4.8-fold increase in the levels of the luciferase mRNA compared with the levels in the pNW1L + 5' Δ 5N-Pac co-transfectants (Fig. 6*B*), indicating that E2F1 can transactivate the *cwp1* promoter.

We further asked whether E2F1 transactivates the *cwp1* promoter through its binding site. The putative E2F1-binding sequence of the *cwp1* promoter was mutated to construct the plasmid pNW1Lm1, and the effect of E2F1 was tested by co-transfection. We found that mutation of the putative E2F1-binding sequence of the *cwp1* promoter resulted in a significant

decrease of the promoter activity to ~30% of the wild type activity during vegetative growth (Fig. 6*A*). In addition, the transactivation levels in the pNW1Lm1 + pPE2F1 co-transfectants decreased significantly (from ~5- to ~2-fold) compared with the pNW1L + 5' Δ 5N-Pac cell line (Fig. 6, *B* and *C*). Therefore, mutation of the putative E2F1-binding sequence in the *cwp1* promoter decreased the transactivation effect by E2F1. The results suggest that E2F1 can transactivate the *cwp1* promoter *in vivo* through its binding site.

The AT-rich Inr element is required for the promoter activity and transcription start site selection (18, 22, 24). Because the

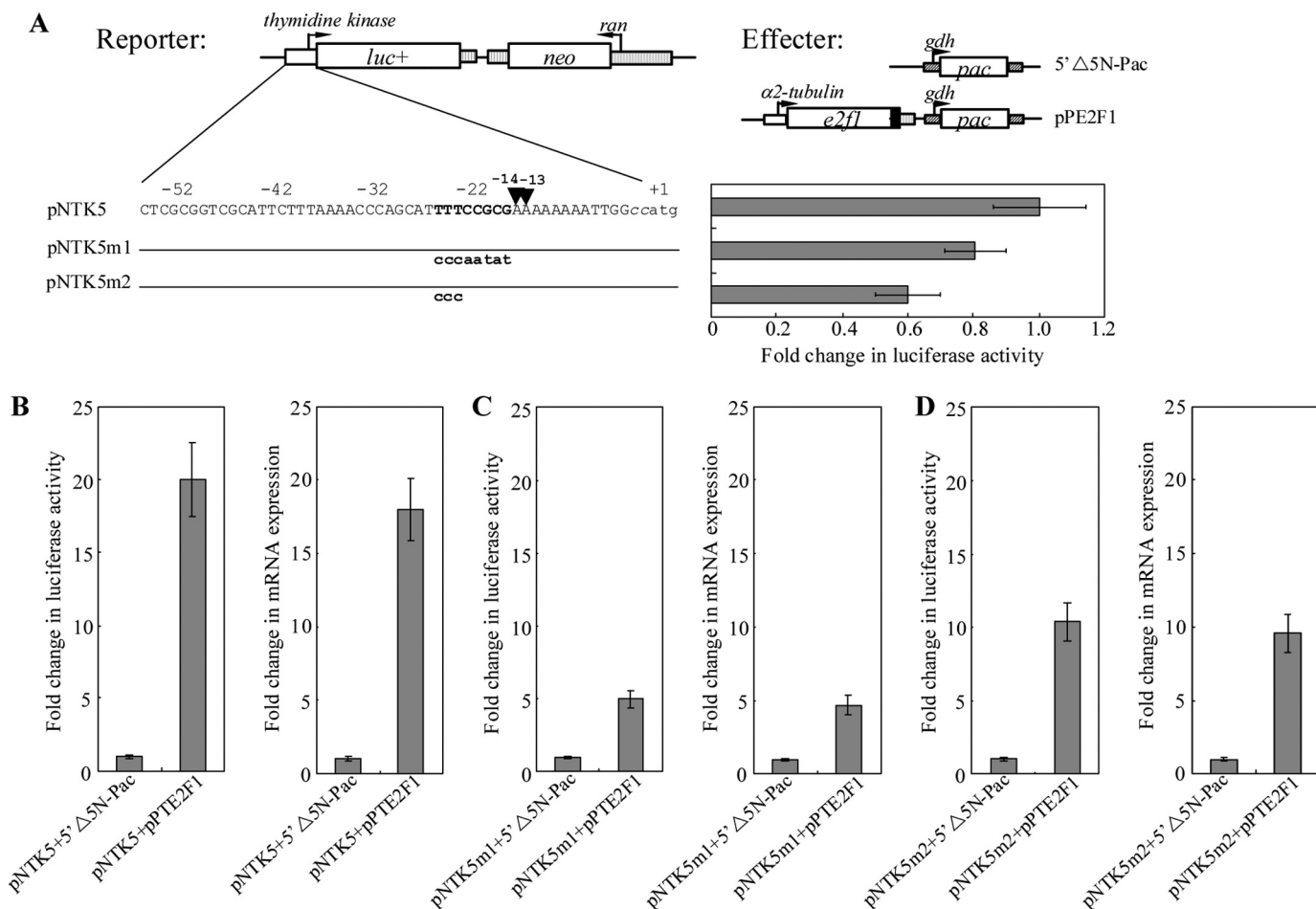


FIGURE 5. Effect of E2F1 on thymidine kinase promoter activity. *A*, diagrams of the pNTK5, pNTK5m1, and pNTK5m2 plasmids. The firefly luciferase gene (*luc+*, open box) is flanked by the 5'-flanking region of the thymidine kinase gene and 3'-flanking region of the *ran* gene (dotted boxes). The *neo* gene is under the control of the 5'- and 3'-flanking regions of the *ran* gene (dotted box). Numbers of the 5'-flanking region of the thymidine kinase gene are relative to the translation start site (+1). Two CC nucleotides inserted upstream of the ATG start codon for plasmid construction are shown in lowercase italics. The putative E2F1 binding sequence is in boldface. The mutated sequence in the construct pNTK5m1 or pNTK5m2 is shown in boldface lowercase type. The transcription start sites determined by 5'-RACE from RNA extracted from vegetative cells are indicated by arrowheads. Diagrams of the 5' Δ5N-Pac and pPE2F1 effector plasmid are the same as in Fig. 2C. Specific cell lines were produced by the stable transfection of reporter construct (pNTK5, pNTK5m1, or pNTK5m2). Luciferase activity was measured in vegetative cells as described under "Experimental Procedures." Fold changes in luciferase expression are shown as the relative expression ratio (experiment/control). Results are expressed as the means ± S.E. of at least three separate experiments (right panel). *B–D*, transactivation of the thymidine kinase promoter by E2F1 in the co-transfection system. Specific cell lines were produced by the stable co-transfection of reporter construct (pNTK5, pNTK5m1, or pNTK5m2) and the effector construct pPE2F1 or the control construct 5' Δ5N-Pac, which expresses only the puromycin selection marker. Luciferase activity was measured in vegetative cells as described under "Experimental Procedures." Real time PCR was performed using primers specific for the luciferase gene and 18 S ribosomal RNA. Transcript levels were normalized to 18 S ribosomal RNA levels. Fold changes in luciferase or mRNA expression are shown as the relative expression ratio (experiment/control). Results are expressed as the means ± S.E. of at least three separate experiments (right panel).

E2F1-binding site is closed to the AT-rich Inr element, it is more likely that there is an interaction between E2F1 and other transcription factors binding to the AT-rich Inr element, such as ARID1 and Pax1 (24, 65). We further asked whether mutation of the AT-rich Inr element affects transactivation function of E2F1. The AT-rich Inr element of the *cwp1* promoter was mutated to construct the plasmid pNW1Lm2, and the effect of E2F1 was tested by co-transfection. We found that mutation of the AT-rich Inr element of the *cwp1* promoter resulted in a significant decrease of the promoter activity to ~1% of the wild type activity during vegetative growth (Fig. 6A). In addition, the transactivation levels in the pNW1Lm2 + pPE2F1 co-transfectants decreased significantly (from ~5- to ~3-fold) compared with the pNW1L + 5' Δ5N-Pac cell line (Fig. 6, B and D). Therefore, mutation of the AT-rich Inr element in the *cwp1* promoter decreased the transactivation effect by E2F1. The results sug-

gest that AT-rich Inr element is important for transactivation by E2F1.

Analysis of E2F1 Mutants—We further identified the portion of E2F1 that is sufficient to interact with DNA or to direct the protein to the nuclei. Two typical nuclear localization signals were predicted in the 11–21 and 230–234 residues of E2F1 using the PSORT program (Fig. 7A) (66). The former nuclear localization signal is located inside the DNA-binding domain 1, and the latter one is located downstream of the two DNA-binding domains (Fig. 7A). We found that mutation of the basic amino acids between residues 11 and 21 (E2F1m1) (Fig. 7A) did not affect nuclear localization in both vegetative and encysting cells (Fig. 7B and supplemental Fig. S1). Mutation of the basic amino acids between residues 230 and 234 (E2F1m2) (Fig. 7A) resulted in a significant decrease of nuclear localization (Fig. 7B and supplemental Fig. S1). The staining was distributed in both

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

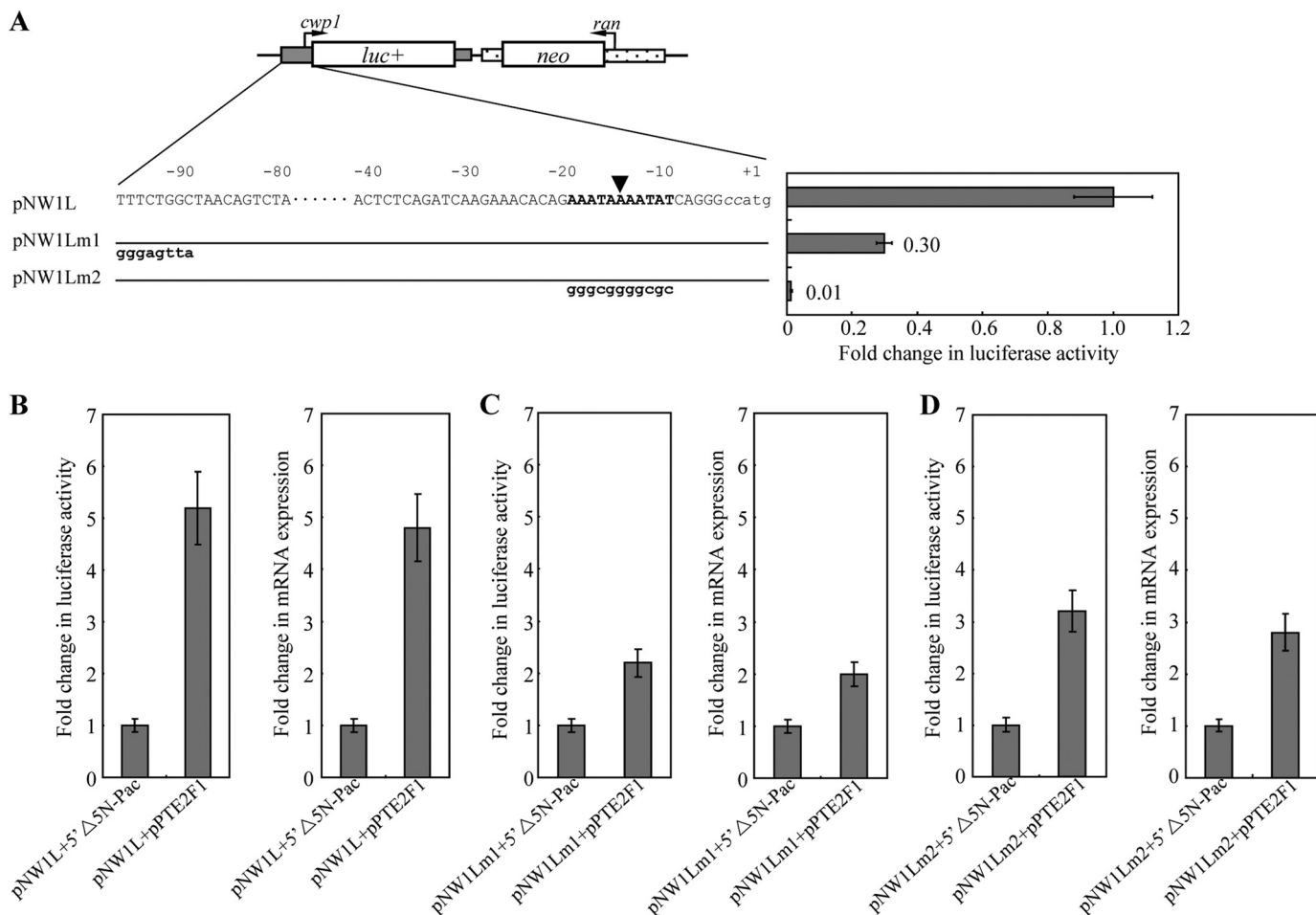


FIGURE 6. Effect of E2F1 on the *cwp1* promoter activity. A, diagrams of the pNW1L, pNW1Lm1, and pNW1Lm2 plasmids. The firefly luciferase gene (*luc+*, open box) is flanked by the 5'-flanking region of the *cwp1* gene and 3'-flanking region of the *ran* gene (dotted boxes). The *neo* gene is under the control of the 5'- and 3'-flanking regions of the *cwp1* gene (dotted box). Numbers of the 5'-flanking region of the *cwp1* gene are relative to the translation start site (+1). Two CC nucleotides inserted upstream of the ATG start codon for plasmid construction are shown in *lowercase italics*. The AT-rich region spanning the transcription start site is in *boldface*. The mutated sequence in the construct pNW1Lm1 or pNW1Lm2 is shown in *boldface lowercase type*. The transcription start site is indicated by an arrowhead (16). Specific cell lines were produced by the stable transfection of reporter construct (pNW1L, pNW1Lm1, or pNW1Lm2). Luciferase activity was measured in vegetative cells as described under "Experimental Procedures." Fold changes in luciferase expression are shown as the relative expression ratio (experiment/control). Results are expressed as the means \pm S.E. of at least three separate experiments (*right panel*). B–D, activation of the *cwp1* promoter by E2F1 in the co-transfection system. Specific cell lines were produced by the stable co-transfection of reporter construct (pNW1L, pNW1Lm1, or pNW1Lm2) and the effector construct pPTE2F1 or the control construct 5' Δ 5N-Pac, which expresses only the puromycin selection marker (Fig. 2C). Luciferase activity was measured in vegetative cells as described under "Experimental Procedures." Real time PCR was performed using primers specific for luciferase gene and 18 S ribosomal RNA. Transcript levels were normalized to 18 S ribosomal RNA levels. Fold changes in luciferase or mRNA expression are shown as the relative expression ratio (experiment/control). Results are expressed as the means \pm S.E. of at least three separate experiments (*right panel*).

the nuclei and cytosol of both vegetative and encysting cells (Fig. 7B and supplemental Fig. S1), suggesting that these basic residues may play an important role in the exclusively nuclear localization. Studies suggest that deletion of the residues making dimerization contact may result in a decrease in the interaction of the human E2F7 and E2F8 and a loss of promoter regulation activity (54). Deletion of the corresponding residues for dimerization contact in giardial E2F1 (residues NVLEA and NILEG) (E2F1dd) (Fig. 7A) resulted in a significant decrease of nuclear localization (Fig. 7B and supplemental Fig. S1). The staining was distributed to both nuclei and some small vesicles in the cytosol (Fig. 7B and supplemental Fig. S1), suggesting that these residues may play an important role in the exclusively nuclear localization. We also determined the role of the conserved RRYXD motif in nuclear localization. Mutation of the second Arg residue in the RRYXD motif of either DNA-binding domain (E2F1DB1m or E2F1DB2m) (Fig. 7A) did not affect

nuclear localization in both vegetative and encysting cells (Fig. 7B and supplemental Fig. S1).

To understand whether the regions we tested for nuclear localization are important for DNA binding, specific E2F1 mutants were expressed in *E. coli*, purified, and tested for their DNA binding activity. Similar levels of wild type E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, and E2F1DB2m were added to the binding reaction mixture (Fig. 7, C and D). We found that mutation of the basic amino acids between residues 11 and 21 (E2F1m1) (Fig. 7A) resulted in a complete loss of binding activity to the tk-30/-1 probe (Fig. 7E). Mutation of the basic amino acids between residues 230 and 234 (E2F1m2) (Fig. 7A) did not change the binding activity to the tk-30/-1 probe (Fig. 7F). We also found that deletion of the corresponding residues for dimerization contact in giardial E2F1 (residues NVLEA and NILEG) (E2F1dd) (Fig. 7A) resulted in a complete loss of binding activity to the tk-30/-1 probe (Fig. 7E). The results suggest

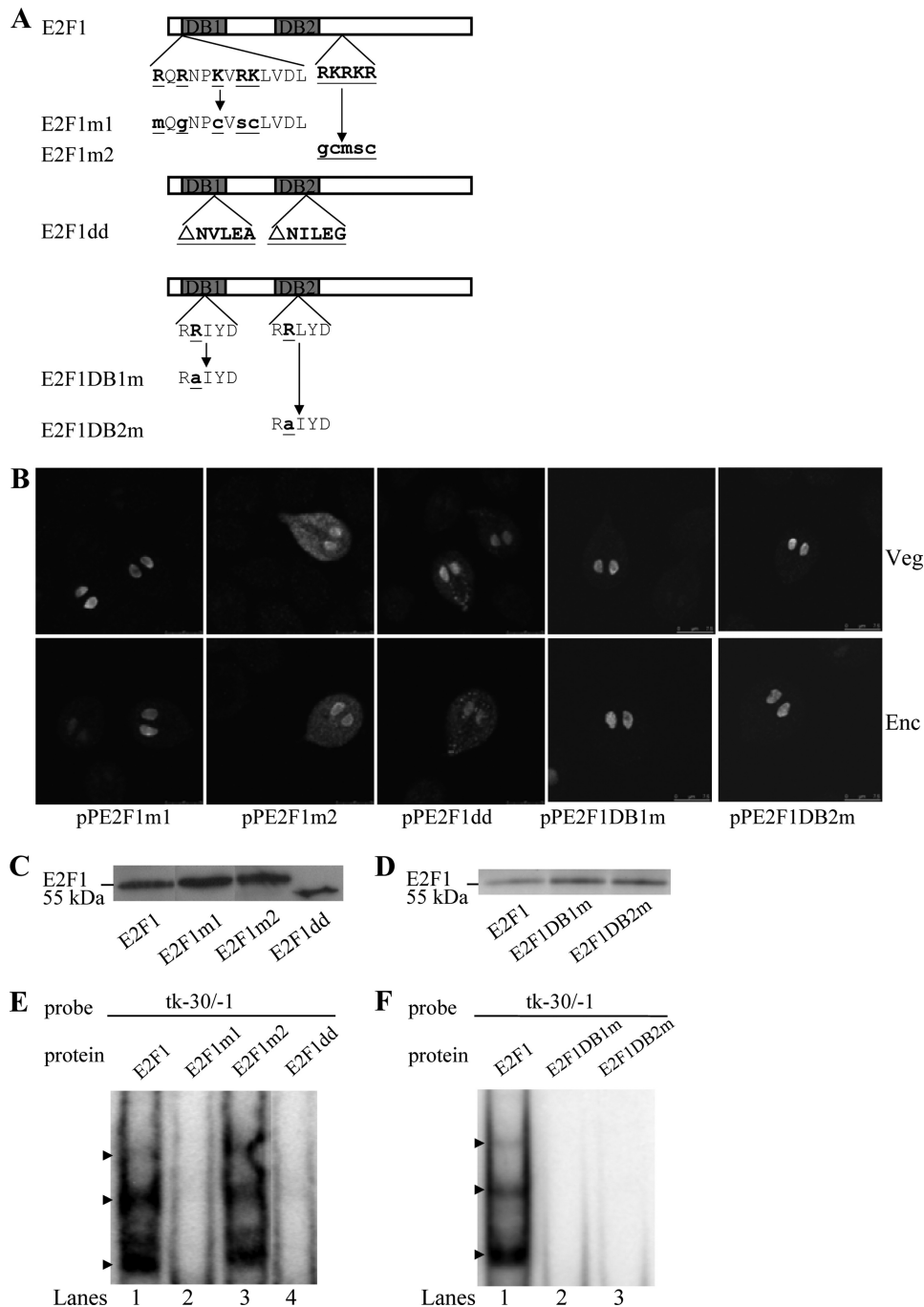


FIGURE 7. Analysis of the DNA-binding domain of E2F1. *A*, diagrams of the E2F1 and mutant proteins. The *gray boxes* indicate the E2F1 DNA-binding domains DB1 and DB2. E2F1m1 (or E2F1m2) contains mutation of stretches of basic amino acids between residues 11 and 23 (or between residues 230 and 234) as shown in *boldface underlined type*. E2F1dd does not contain NVLEA (residues 61–65) and NILEG (residues 155–159) sequence. In E2F1DB1m and E2F1DB2m, the second Arg of the RRYXD DNA recognition motifs are mutated as shown in *underlined type*. *B*, localization of the E2F1 mutants. The *e2f1* gene was mutated as described (*A*) and subcloned to replace the wild type *e2f1* gene in the backbone of pPE2F1 (Fig. 2C); the resulting plasmids pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, and pPE2F1DB2m were transfected into *Giardia*. The stable transfectants were cultured in growth (*Veg*, vegetative growth; *upper panels*) or encystation medium for 24 h (*Enc*, encystation; *lower panels*) and then subjected to immunofluorescence analysis using anti-HA antibody for detection. The products of pPE2F1m1, pPE2F1DB1m, and pPE2F1DB2m localize to the nuclei in both vegetative and encysting trophozoites. The product of pPE2F1m2 and pPE2F1dd localize not only in nucleus, suggesting their partial loss of nuclear localization. The product of pPE2F1m2 localizes to the nuclei and cytosol in both vegetative and encysting trophozoites. The product of pPE2F1dd localizes to the nuclei and some vesicles in cytosol in both vegetative and encysting trophozoites. *C* and *D*, Western blot analysis of recombinant E2F1 and mutant proteins. The E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m protein with a V5 tag at its C terminus was purified by affinity chromatography and then detected by anti-V5-HRP antibody in Western blots. *E* and *F*, reduction of DNA binding ability of E2F1 mutants. Electrophoretic mobility shift assays were performed using purified E2F1, E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, E2F1DB2m, and tk-30/-1 probe. The *arrowheads* indicate the shifted complexes. A positive control reaction was carried out using the wild type E2F1 (*lane 1*).

that the residues inside the DNA-binding and dimerization domain may play an important role in DNA binding. Each RRYXD motif is known to be important for DNA binding of

E2Fs with two DNA-binding domains (43, 55). We found that E2F1DB1m or E2F1DB2m with a mutation of the second Arg residue in the RRYXD motif (Fig. 7A) did not bind to the

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

tk-30/-1 probe (Fig. 7F), indicating that the RXYD motif is important for DNA binding.

Overexpression of E2F1 Induced the Expression of *cwp1-3*, and *myb2* Genes—To study the role of E2F1 in *G. lamblia*, we expressed *e2f1* by the α 2-tubulin promoter (pPE2F1; Fig. 2C) and observed its gene expression. An ~55-kDa protein was detected (Fig. 8A), which was almost matched to the predicted molecular mass of E2F1 (~46.2 kDa) with the HA tag (~1 kDa). Similar to the expression pattern of the endogenous E2F1 protein, the levels of the E2F1-HA protein increased significantly during encystation (data not shown; also see Fig. 2B). Overexpression of E2F1 in the pPE2F1 cell line can be confirmed by the anti-E2F1 antibody, although the endogenous E2F1 and overexpressed E2F1-HA migrated at similar sizes (Fig. 8A). We found that E2F1 overexpression resulted in a significant increase of the CWP1 protein levels during vegetative growth (Fig. 8A). RT-PCR and quantitative real time PCR analysis showed that the mRNA levels of the endogenous *e2f1* plus vector expressed *e2f1* in the E2F1-overexpressing cell line increased by ~3.7-fold ($p < 0.05$) (Fig. 8, B and C) relative to the control cell line, which expressed only the puromycin selection marker (5'Δ5N-Pac) (Fig. 2C) (60). The mRNA levels of the endogenous *cwp1*, *cwp2*, *cwp3*, *myb2*, and thymidine kinase genes in the E2F1-overexpressing cell line increased by ~2.3–2.6-fold ($p < 0.05$) relative to the control cell line (Fig. 8, B and C, data not shown). Similar mRNA levels of the *ran* and 18 S ribosomal RNA genes were detected (Fig. 8, B and C). We further investigated the effect of giardial E2F1 on cyst formation. In previous studies, we have found that some *G. lamblia* trophozoites may undergo spontaneous differentiation (57). We obtained consistent cyst counts data for vegetative *G. lamblia* cultures during growth to stationary phase (~4800 cysts/ml for 5'Δ5N-Pac cell line) (60). In this study, we found that the cyst number in the E2F1-overexpressing cell line increased by ~2.3-fold ($p < 0.05$) relative to the control cell line, which expresses only the puromycin selection marker (5'Δ5N-Pac) (Fig. 2C), indicating that the overexpressed E2F1 can increase the cyst formation (Fig. 8D). Similar results were obtained during encystation (data not shown). The results suggest that the overexpressed E2F1 can transactivate the *cwp1*, *cwp2*, *cwp3*, *myb2*, and thymidine kinase genes.

To further understand the function of giardial E2F1, we observed the effect of overexpression of the E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, and E2F1DB2m (Fig. 8). E2F1m1, E2F1DB1m, and E2F1DB2m can enter nuclei, but E2F1m2 and E2F1dd partially localized to nuclei (Fig. 7B and [supplemental Fig. S1](#)). We found that the levels of E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m protein increased significantly compared with that of wild type E2F1 during vegetative growth in both anti-HA and anti-E2F1 Western blots (Fig. 8A). We further analyzed whether the transcript levels of the E2F1m1, E2F1m2, E2F1dd, E2F1DB1m, or E2F1DB2m were changed. As shown by RT-PCR and quantitative real time PCR analysis, the levels of HA-tagged *e2f1m1*, *e2f1m2*, *e2f1dd*, *e2f1db1m*, or *e2f1db2m* mRNA increased by ~1.3–3.1-fold ($p < 0.05$) compared with that of wild type HA-tagged *e2f1* during vegetative growth (Fig. 8, B and C). We did not detect any HA-tagged *e2f1* transcripts in the 5'Δ5N-Pac control cell

line (Fig. 8B). We also found that the levels of the CWP1 protein and the *cwp1*, *cwp2*, *cwp3*, *myb2*, and thymidine kinase mRNA and cyst formation increased significantly in the E2F1m1-, E2F1dd-, E2F1DB1m-, or E2F1DB2m-overexpressing cell line relative to the wild type E2F1-overexpressing cell line (Fig. 8, A–D). Similar results were obtained during encystation (data not shown). The results suggest an increase of transactivation activity of E2F1m1, E2F1dd, E2F1DB1m, and E2F1DB2m whose expression also increased significantly. Therefore, it is possible that these E2F1 mutants still possess transactivation activity. However, the transactivation activity of the E2F1m2 mutant did not increase (Fig. 8, A–D), possibly because of its partial mis-localization and its relatively lower expression as compared with other E2F1 mutants (Figs. 7B and 8C).

Oligonucleotide microarray assays confirmed the up-regulation of the *cwp1*, *cwp2*, *cwp3*, *myb2*, and thymidine kinase gene expression in the E2F1-overexpressing cell line to ~2.12 to ~2.7-fold of the levels in the control cell line (Fig. 8E). Similar mRNA levels of the *ran* gene were detected (Fig. 8E). We found that 31 and 42 genes were significantly up-regulated (>2-fold) and down-regulated (<1/2) ($p < 0.05$) in the E2F1 overexpression cells relative to the vector control, respectively (Fig. 8E and Table 1). The expression levels of the *e2f1* gene in the E2F1-overexpressing cell line increased by ~3.65-fold ($p < 0.05$) (Fig. 8E and Table 1).

Interaction between E2F1 and Myb2—It is possible that E2F1 may regulate target genes by interacting with other transcription factors. We further tried to understand whether E2F1 can interact with Myb2, which is an encystation-induced transcription factor and is involved in coordinate up-regulation of the *cwp1-3* genes during encystation (21, 26). We used glutathione S-transferase (GST) pulldown assays to investigate the possibility that E2F1 interacts with Myb2. We performed the assays using the GSTMyb2 fusion protein bound in the glutathione-Sepharose beads and purified E2F1. Specific pulldown interactions were observed with GSTMyb2 and E2F1, whereas no interaction was observed with GST and E2F1 (Fig. 9, A–C).

We then performed co-immunoprecipitation experiments using the E2F1-overexpressing cell line. We lysed the cells and immunoprecipitated the HA-tagged E2F1 with anti-HA antibody. Western blots of immunoprecipitates probed with anti-HA and anti-Myb2 indicate that E2F1 co-precipitates with Myb2 (Fig. 9, D and E). As a control, the anti-HA antibody did not immunoprecipitate E2F1 and Myb2 in the control cell line, which expressed only the puromycin selection marker (5'Δ5N-Pac) (Fig. 2C) (60), suggesting that Myb2 co-immunoprecipitated with anti-HA requires the HA-tagged E2F1 protein (Fig. 9, D and E). The results suggest an interaction between E2F1 and Myb2.

DISCUSSION

The E2F protein family is a group of transcription factors that regulate cell cycle progression and cell differentiation in a wide variety of species, including *Caenorhabditis elegans*, *Dictyostelium discoideum*, plants, *Drosophila*, and mammals (42, 68, 69). It has not been identified to date in yeast, but an E2F functional analog, SWI4, is present in yeast (70, 71). SWI4 has a different type of DNA-binding domain as compared with the E2F DNA-binding domain, although both SWI4 and E2F bind to similar target sequences (70). In this study, an E2F-like transcription

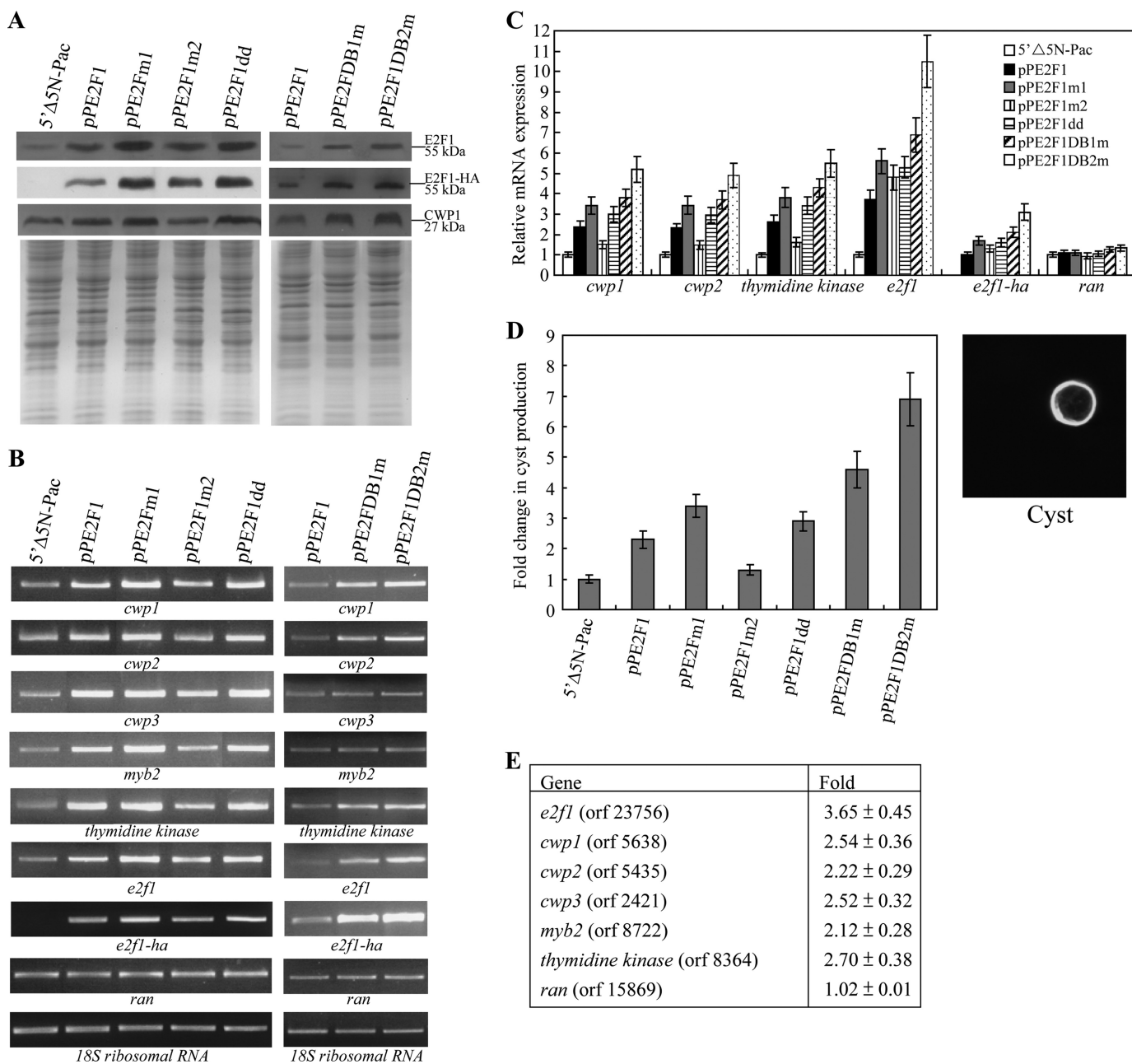


FIGURE 8. Activation of *cwp1*–*3*, *myb2*, and *thymidine kinase* gene expression in the E2F1-overexpressing cell line. *A*, overexpression of E2F1 increased the levels of CWP1 protein. The 5'Δ5N-Pac, pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, and pPE2F1DB2m stable transfectants were cultured in growth medium and then subjected to SDS-PAGE and Western blot. The blot was probed by anti-E2F1, anti-HA, and anti-CWP1 antibody. Equal amounts of protein loading were confirmed by SDS-PAGE and Coomassie Blue staining. Representative results are shown. *B*, RT-PCR analysis of gene expression in the E2F1-overexpressing and E2F1 mutants overexpressing cell lines. The 5'Δ5N-Pac, pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, and pPE2F1DB2m stable transfectants were cultured in growth medium and then subjected to RT-PCR analysis. PCR was performed using primers specific for *e2f1-ha*, *e2f1*, *cwp-3*, *myb2*, *thymidine kinase*, *ran*, and 18 S ribosomal RNA genes. *C*, quantitative real time PCR analysis of gene expression in the E2F1-overexpressing and E2F1 mutants overexpressing cell lines. The 5'Δ5N-Pac, pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, and pPE2F1DB2m stable transfectants were cultured in growth medium and then subjected to quantitative real time PCR analysis. Real time PCR was performed using primers specific for *e2f1-ha*, *e2f1*, *cwp1*, *cwp2*, *thymidine kinase*, *ran*, and 18 S ribosomal RNA genes. Similar mRNA levels of the *ran* and 18 S ribosomal RNA genes for these samples were detected. Transcript levels were normalized to 18 S ribosomal RNA levels. Fold changes in mRNA expression are shown as the ratio of transcript levels in the pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, or pPE2F1DB2m cell lines relative to the 5'Δ5N-Pac cell line. Results are expressed as the means ± S.E. of at least three separate experiments. *D*, cyst count. The 5'Δ5N-Pac, pPE2F1, pPE2F1m1, pPE2F1m2, pPE2F1dd, pPE2F1DB1m, and pPE2F1DB2m stable transfectants were cultured in growth medium and then subjected to cyst count as described under "Experimental Procedures." The sum of total cysts is expressed as relative expression level over control. Values are shown as means ± S.E. (left panel). Cysts were subjected to immunofluorescence analysis, using anti-CWP1 antibody for detection. The CWP1 localizes to the cyst wall in a representative cyst (right panel). *E*, microarray analysis. Microarray data were obtained from the 5'Δ5N-Pac and pPE2F1 cell lines during vegetative growth. Fold change are shown as the ratio of transcript levels in the pPE2F1 cell line relative to the 5'Δ5N-Pac cell line. Results are expressed as the means ± S.E. of at least three separate experiments.

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

TABLE 1
Genes up or down regulated by E2F1 overexpression

No.	Annotation	Orf no.	Fold change (pPE2F1/5'Δ5N-Pac)
1	Hypothetical protein	112080	8.31 ($p < 0.05$) ^a
2	High cysteine membrane protein group 4	114930	7.76 ($p < 0.05$)
3	Hypothetical protein	117393	6.90 ($p < 0.05$)
4	VSP	14331	4.84 ($p < 0.05$)
5	High cysteine membrane protein group 1	25816	3.88 ($p < 0.05$)
6	E2F1	23756	3.65 ($p < 0.05$)
7	Hypothetical protein	32657	3.57 ($p < 0.05$)
8	Chorein	87358	3.48 ($p < 0.05$)
9	Hypothetical protein	9068	3.45 ($p < 0.05$)
10	Hypothetical protein	101484	3.38 ($p < 0.05$)
11	VSP	112113	3.32 ($p < 0.05$)
12	High cysteine membrane protein VSP-like	112305	3.21 ($p < 0.05$)
13	Hypothetical protein	112696	2.97 ($p < 0.05$)
14	Hypothetical protein	113433	2.96 ($p < 0.05$)
15	Thymidine kinase	8364	2.70 ($p < 0.05$)
16	High cysteine membrane protein group 4	114089	2.63 ($p < 0.05$)
17	DNA repair protein RAD51	13104	2.62 ($p < 0.05$)
18	VSP	137620	2.55 ($p < 0.05$)
19	Cyst wall protein 1	5638	2.54 ($p < 0.05$)
20	Cyst wall protein 3	2421	2.52 ($p < 0.05$)
21	VSP	14307	2.43 ($p < 0.05$)
22	VSP	26590	2.37 ($p < 0.05$)
23	Hypothetical protein	28388	2.37 ($p < 0.05$)
24	Hypothetical protein	31420	2.37 ($p < 0.05$)
25	VSP	40630	2.36 ($p < 0.05$)
26	VSP	41476	2.34 ($p < 0.05$)
27	Hypothetical protein	9551	2.33 ($p < 0.05$)
28	Hypothetical protein	97698	2.33 ($p < 0.05$)
29	Cyst wall protein 2	5435	2.22 ($p < 0.05$)
30	Myb2	8722	2.12 ($p < 0.05$)
31	Hypothetical protein	11148	2.12 ($p < 0.05$)
32	VSP	101765	0.11 ($p < 0.05$)
33	Clathrin heavy chain	102108	0.20 ($p < 0.05$)
34	High cysteine protein	102180	0.20 ($p < 0.05$)
35	VSP with INR	101074	0.22 ($p < 0.05$)
36	Variant-specific surface protein	103237	0.23 ($p < 0.05$)
37	High cysteine protein	112604	0.27 ($p < 0.05$)
38	VSP	112678	0.27 ($p < 0.05$)
39	VSP	113269	0.27 ($p < 0.05$)
40	VSP with INR	113450	0.29 ($p < 0.05$)
41	Proprotein convertase subtilisin/kexin type 5 precursor	114625	0.31 ($p < 0.05$)
42	Dynein heavy chain	101138	0.31 ($p < 0.05$)
43	VSP	114672	0.32 ($p < 0.05$)
44	Hypothetical protein	114674	0.32 ($p < 0.05$)
45	VSP	115796	0.32 ($p < 0.05$)
46	Ceramide glucosyltransferase	11642	0.33 ($p < 0.05$)
47	Hypothetical protein	119703	0.34 ($p < 0.05$)
48	VSP	122566	0.35 ($p < 0.05$)
49	Hypothetical protein	123336	0.35 ($p < 0.05$)
50	VSP AS8	13194	0.36 ($p < 0.05$)
51	VSP	13390	0.36 ($p < 0.05$)
52	VSP	137606	0.36 ($p < 0.05$)
53	VSP	137612	0.36 ($p < 0.05$)
54	VSP	137617	0.36 ($p < 0.05$)
55	VSP with INR	14586	0.37 ($p < 0.05$)
56	Hypothetical protein	14690	0.38 ($p < 0.05$)
57	Hypothetical protein	15419	0.39 ($p < 0.05$)
58	Coiled-coil protein	16199	0.39 ($p < 0.05$)
59	Liver stage antigen-like protein	16595	0.39 ($p < 0.05$)
60	Kinase, NEK	17084	0.40 ($p < 0.05$)
61	Hypothetical protein	17332	0.41 ($p < 0.05$)
62	CXC-rich protein	17476	0.41 ($p < 0.05$)
63	VSP	34357	0.42 ($p < 0.05$)
64	Hypothetical protein	36122	0.42 ($p < 0.05$)
65	Hypothetical protein	3731	0.42 ($p < 0.05$)
66	VSP	40591	0.42 ($p < 0.05$)
67	Hypothetical protein	5206	0.44 ($p < 0.05$)
68	Variant-specific surface protein	6101	0.45 ($p < 0.05$)
69	High cysteine membrane protein group 1	7715	0.45 ($p < 0.05$)
70	Hypothetical protein	93278	0.46 ($p < 0.05$)
71	Coiled-coil protein	9515	0.47 ($p < 0.05$)
72	Hypothetical protein	95908	0.49 ($p < 0.05$)
73	Hypothetical protein	99726	0.49 ($p < 0.05$)

^a p values were determined for groups in which the average means changed by a factor of ≥ 2.0 or ≤ 0.5 .

factor has been identified in *G. lamblia*, although it is divergent in sequence. This suggests that the E2F family may have evolved before divergence of *G. lamblia* from the main eukaryotic line

of descent but may have been lost in yeast. To date, giardial E2F is the first E2F transcription factor identified in early diverging protozoan parasites. We did BLAST searches for the genome

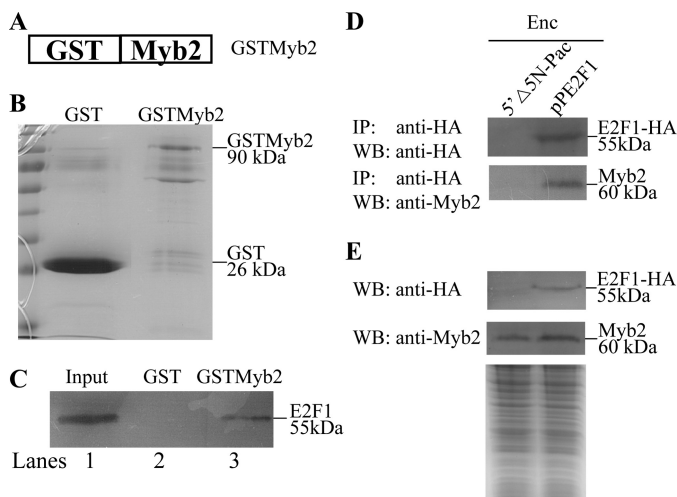


FIGURE 9. Interaction between E2F1 and Myb2. *A*, schematic representation of GSTMyb2 fusion protein. Full-length Myb2 was fused to the C terminus of GST. *B*, presence of bound GST and GSTMyb2 in the glutathione-Sepharose beads. Following binding of GST and GSTMyb2 and extensive washing, the glutathione-Sepharose beads were subjected for SDS-PAGE. GST and GSTMyb2 were visualized by Coomassie Blue staining. *C*, interaction between E2F1 and Myb2 by GST pull-down assay. Purified recombinant E2F1 with a C-terminal V5 tag was mixed with GST protein, and GSTMyb2 fusion proteins in the glutathione-Sepharose beads and the pull-down fractions were analyzed. Five percent of the input (*lane 1*) and 25% of the pull-down material (*lanes 2 and 3*) were subjected to Western blot analysis. V5-tagged E2F1 was detected using anti-V5-HRP antibody. *D*, co-immunoprecipitation assays. The 5' Δ5N-Pac and pPE2F1 stable transfectants were immunoprecipitated (IP) using anti-HA antibody conjugated to beads. The precipitates were analyzed by Western blot (WB) with anti-HA or anti-Myb2 antibody as indicated. *E*, expression of the HA-tagged E2F1 and Myb2 proteins in whole cell extracts. The 5' Δ5N-Pac and pPE2F1 stable transfectants were cultured in encystation medium for 24 h (Enc, encystation) and then subjected to Western blot analysis. The blot was probed by anti-HA and anti-Myb2 antibody. Equal amounts of protein loading were confirmed by SDS-PAGE and Coomassie Blue staining.

databases of *Entamoeba histolytica*, *Plasmodium falciparum*, and *Trypanosoma brucei* and only identified matches for uncharacterized proteins. At least two putative *e2f*-like genes were present in the *Trichomonas vaginalis* genome database. Further analyses will reveal whether E2F family proteins are shared with these eukaryotic lineages.

The key components of the giardial cyst wall, cyst wall proteins, are synthesized during *Giardia* differentiation into dormant cysts (2). E2F proteins in higher eukaryotes are involved in cell differentiation and function as transcriptional activators or repressors (42, 43, 49, 54, 55). It has been found that eukaryotic E2F may regulate S phase-specific genes, including thymidine kinase, of which product is a key enzyme of DNA synthesis (31–33, 35). To gain insight into the function of E2F1 in *Giardia*, we tested the hypothesis that E2F1 can regulate the thymidine kinase gene. Our results show that the giardial E2F1 localizes to the cell nuclei (Fig. 2D and supplemental Fig. S1). The increased levels of the E2F1 protein during encystation indicate that it may also play a role in this differentiation. We also found an increase of the thymidine kinase gene expression during encystation (Fig. 2A). This is possibly due to the requirement of DNA replication during encystation. E2F1 can bind to a specific sequence in the core promoter region of the thymidine kinase gene (Fig. 3B). Mutation of the E2F1-binding site in the thymidine kinase promoter resulted in a decrease of promoter activity (Fig. 5), suggesting that E2F1 may be a transcrip-

tional activator in the regulation of the thymidine kinase gene. We further analyzed the function of E2F1 in *Giardia* encystation. We found that E2F1 can bind to specific sequences in the core promoter region of the *cwp1*, *cwp2*, and *cwp3* genes (Fig. 3B). Mutation of the E2F1-binding site in the *cwp1* promoter resulted in a decrease of promoter activity (Fig. 6), suggesting that E2F1 may be a transcriptional activator in the regulation of the *cwp1* gene. Interestingly, the constitutively overexpressed E2F1 increased the levels of the *cwp1–3*, *myb2*, and thymidine kinase mRNA (Fig. 8, B and C). The levels of the CWP1 protein and cyst formation also increased in the E2F1-overexpressing cell line (Fig. 8, A and D). ChIP assays also confirmed the association of E2F1 with its own promoter and the *cwp1–3*, *myb2*, and thymidine kinase promoters (Fig. 4C). We also found an important role of the basic residues of the E2F1 DNA-binding domain in the exclusive nuclear localization (Fig. 7B and supplemental Fig. S1). The results suggest that E2F1 may play an important role in induction of encystation-induced *cwp* and thymidine kinase genes.

The giardial promoters defined to date, including the encystation-specific *cwp* promoters, are notably short and contain AT-rich Inr elements (15, 16, 18, 19–23). Deletion and mutation analysis of the *ran*, α 2-tubulin, and *cwp2* promoters has provided the evidence that the AT-rich Inr elements are positive *cis*-acting elements and that they are important for basal promoter activity and transcription start site selection (18, 19, 22, 24). Previously, we have identified several transcription factors whose expression increased significantly during encystation and whose function may be related to the transactivation of the *cwp* genes, including Myb2, GARP-like protein 1, ARID1, WRKY, and Pax1 (21, 24–27, 65). ARID1 and Pax1 can bind to the AT-rich Inr elements of the *cwp* promoters (24, 65). Myb2, GARP-like protein 1, and WRKY can bind to the proximal upstream regions of the *cwp* promoters, and their binding sequences are positive *cis*-acting elements (21, 25, 27). In this study, we found that E2F1 can also bind to specific sequences upstream of the AT-rich Inr elements. There may be an interaction of the transcription factors binding to the proximal upstream regions and the AT-rich Inr elements. This interaction may be required for promoter activity and accurate transcription start site selection.

Our results also showed that E2F1 can bind weakly to the *cwp1–104/–70* probe *in vitro*, which contains the sequence TTTCTGGC (Fig. 3B). This sequence is similar to the E2F1-binding sequence TTTCCGCG in the thymidine kinase promoter (Fig. 3B), suggesting that E2F1 can bind to the *cwp1* promoter, although weakly. We also detected the association of E2F1 with the *cwp1* promoter *in vivo* (Fig. 4C), suggesting that E2F1 may activate the *cwp1* promoter directly. We also found the presence of the putative E2F1-binding site in the *cwp2* and *cwp3* promoters and an association of E2F1 with the *cwp2* and *cwp3* promoters *in vivo* (Fig. 3B), suggesting that E2F1 may activate the *cwp2* and *cwp3* promoters directly. Although E2F1 can also function as a transactivator, it may still need to cooperate with other transcription factors that are induced during encystation to transactivate the *cwp* genes. We found that mutation of the E2F1 target sequences in the thymidine kinase promoter might impair the binding of E2F1, leading to a significant decrease in transactivation (Figs. 4 and 5). Mutation of the

Regulation of *cwp* Genes by E2F1 in *Giardia lamblia*

E2F1 target sequences or the AT-rich Inr in the *cwp1* promoter might impair the binding of E2F1 or other Inr-binding proteins to the Inr, leading to a significant decrease in transactivation (Fig. 6). The results suggest that E2F1 may transactivate the *cwp1* promoter *in vivo* through its DNA-binding sequences, and there may be an interaction of the E2F1 or other Inr-binding transcription factors. This interaction may be required for transcriptional activation of the *cwp1* gene during encystation.

The giardial E2F1 contains two DNA-binding domains near the N-terminal region, similar to the human E2F7–8 and plant E2L1–3 (30, 43, 53–55). Although divergent from human E2F proteins, giardial E2F1 binds DNA in a sequence-specific manner. Our results show that the giardial E2F1-binding site TTTCCGCG is similar to the human E2F-binding sites TTT(c/g)GCGC(c/g) or TTTCCCGCC (42, 45–52). This indicates that the E2F family-binding sites may have been conserved in evolution. This may reflect similar function of the E2F family. From the results with the human E2F4 (49), the most important residues contacting or helping contact the bases G and C of the two strands of binding sequence (TTTCGCGCG) are two Args, one Tyr, and one Asp of RRXYD motif of the α 3 helix. These residues are conserved in the two DNA-binding domains of the giardial E2F1. We also found that the second Arg of RRXYD motif was also required for giardial E2F1 binding (see below), indicating the importance of this residue. The Arg of the α N helix contacting the T-rich region of the binding sequence is not conserved in the two DNA-binding domains of the giardial E2F1 (Fig. 1B). The giardial E2F1 has some but not all of the conserved residues that make DNA backbone contact or dimerization contact (Fig. 1B). These nonconservative substitutions might explain why giardial E2F1 can recognize longer target sequences, including the AT-rich Inr sequence located downstream of the TTTCCGCG sequence.

Two putative nuclear localization signals have been found, of which the one downstream of the DNA-binding domain (residues 230–234) but not the one inside the DNA-binding domain (residues 11–21) is important for nuclear localization (Fig. 7B and supplemental Fig. S1). However, the one inside the DNA-binding domain (residues 11–21) but not the one downstream of the DNA-binding domain (residues 230–234) is important for DNA binding (Fig. 7C). We also found that deletion of the corresponding residues for dimerization contact in giardial E2F1 (residues NVLEA and NILEG were deleted in E2F1dd) not only resulted in a significant decrease of nuclear localization, but also resulted in a complete loss of DNA binding activity (Fig. 7, B and C). The results suggest that the residues inside the DNA-binding domain and dimerization domain may play an important role in DNA binding. Each RRXYD motif is known to be important for DNA binding of E2Fs with two DNA-binding domains (43, 55). We found that mutation of the second Arg residue in the RRXYD motif (E2F1DB1m or E2F1DB2m) did not affect nuclear localization but resulted in a loss of DNA binding activity (Fig. 7, B and D), indicating that the RRXYD motif is important for DNA binding.

pRB family proteins bind the transactivation domain of E2F1–3 transactivators to inhibit the transactivation function in human (37–42). Sequences for binding to pRB family proteins are also present in the E2F4 and E2F5 repressors but not

present in E2F6–8 repressors (42). The C-terminal end of the giardial E2F1 is not highly similar to the C-terminal pRB-binding domain of the human E2Fs (Fig. 1C). We did not find consensus pRB-binding sequences in the giardial E2F1 protein. We did blast searches of the *G. lamblia* genome database but did not identify matches for the pRB family proteins (data not shown). This suggests that the pRB family proteins may be too divergent to be recognized by sequence similarity. It is also possible that they are not present in *G. lamblia*, but their functional analogs are present in *G. lamblia*. This awaits further studies to explore. pRB has not been identified in yeast, although its functional analog, Whi5, is present (71).

Many important transcription factors involved in developmental regulation and in stress response have an autoregulation mechanism, including mammalian c-Myb and Pax proteins and plant WRKY (72–74). Myb2, Pax1, or WRKY has been found to be positively or negatively autoregulated to maintain its own protein levels in *G. lamblia*, and this is related to the presence of its binding sites in its own promoter region (21, 26, 27, 65). It has been shown that mammalian E2F1 proteins may be negatively autoregulated by an E2F·DP-pRB complex in quiescent cells and that they may be positively autoregulated by activating the activity of its own promoter to maintain high levels of E2F1 protein during G₁/S (75). We found that some E2F1 mutants, including E2F1m1, E2F1dd, E2F1DB1m, and E2F1DB2m, have an increase of transactivation function on the *cwp* and thymidine kinase gene promoters (Fig. 8, B and C). Their transactivation function was correlated with their increased levels. Their mRNA levels increased significantly compared with that of wild type E2F1, suggesting a positive autoregulation of the *e2f1* gene. Transactivation function of specific transcription factors may require interaction with transcriptional co-activators and general transcription factors. It is possible that the E2F1m1, E2F1dd, E2F1DB1m, and E2F1DB2m mutants may possess minor mutations on the DNA-binding domains. They may still possess sufficient transactivation activity and ability to interact with other transcriptional co-activators. However, the transactivation activity of the E2F1m2 mutant did not increase, possibly because of its partial mis-localization and its relatively lower expression as compared with other E2F1 mutants. It is also possible that the E2F1m2 mutant may possess some mutations downstream of the DNA-binding domains, and the region of mutations may be located in a transactivation domain. The mutations may decrease its transactivation activity and ability to interact with other transcriptional co-activators.

We found that E2F1 cannot bind to the 18 S ribosomal RNA gene promoter that does not contain the putative E2F1-binding site in the <200-bp 5'-flanking region (Fig. 4C). E2F1 can bind to the promoters of the constitutive *ran* gene and the encystation-induced *cwp1–3*, *myb2*, and thymidine kinase genes *in vivo* (Fig. 4C), suggesting that E2F1 may be involved in transcriptional regulation of many different genes. However, overexpressed E2F1 did not transactivate the *ran* gene promoter (Fig. 8, B and C). Oligonucleotide microarray assays confirmed the up-regulation of the *cwp1*, *cwp2*, *cwp3*, *myb2*, and thymidine kinase gene expression in the E2F1-overexpressing cell line, respectively (Fig. 8E). We also found that 31 and 42 genes were significantly up-regulated (>2-fold) and down-regulated

(<1/2) ($p < 0.05$) in the E2F1 overexpression cells relative to the vector control (Fig. 8E and Table 1). Interestingly, DNA repair protein RAD51 gene was up-regulated by E2F1 (Fig. 8E and Table 1). It has been shown that the expression of human RAD51 is induced in late S phase and G₂/M phase (76). E2F proteins can regulate DNA repair protein genes, including the RAD51 gene to integrate cell cycle progression (77). Our results suggest that the giardial E2F1 may have a similar function in the regulation of DNA repair protein RAD51 gene.

In late-branching eukaryotes, E2F proteins regulate specific target genes by interacting with other classes of transcription factors, including Sp1, P53, BRCA, NF- κ B, AT-rich interaction domain, and YY1 (78–84). Therefore, it is possible that giardial E2F1 functions as an activator via association with some encystation-specific cofactors on the promoter context of encystation-induced genes. In late-branching eukaryotes, E2Fs and B-Myb have an important role in the control of G₁/S and G₂/M genes, respectively (85). The *b-myb* gene is also a target for E2Fs at G₁/S (85). Interestingly, there is an interaction between E2Fs and B-Myb, and binding of B-Myb to the G₂-specific *cdc2* (*cdk1*) gene promoter is dependent on an intact E2F-binding site, suggesting that E2Fs may link the G₁/S and G₂/M transcriptional regulation by interaction with B-Myb (85). We also found an interaction between giardial E2F1 and Myb2 using *in vitro* GST pulldown assays and *in vivo* co-immunoprecipitation assays. The presence of the Myb2-binding sites in the promoters of key encystation-induced genes *cwp1-3*, *g6pi-b*, and *myb2* itself suggests that Myb2 may be involved in coordinating their differential expression (21, 26). CHIP assays have been used to confirm the binding of Myb2 and E2F1 to the *cwp1*, *cwp2*, and *myb2* gene promoters (Fig. 4C) (26). The interaction of E2F1 with Myb2 suggests that E2F1 may participate in activating expression of the *cwp* genes during giardial encystation and that E2F1 and Myb2 may have a synergistic effect on activated transcription of *cwp* genes during encystation.

Our results indicate that E2F1 can transactivate the thymidine kinase and *cwp* genes that are involved in DNA synthesis and differentiation in the primitive protozoan *G. lamblia*. Our study provides evidence for the important role of the giardial E2F1 in the differentiation of *G. lamblia* trophozoites into cysts, leading to greater understanding of the evolution of eukaryotic DNA-binding domain and transcriptional mechanisms during cell differentiation.

Acknowledgments—We thank Dr. Chien-Kuo Lee, Dr. Tsai-Kun Li, Dr. Shin-Hong Shiao, and Dr. Nei-Li Chan for helpful comments; Yi-Li Liu and I-Ching Huang for technical support in DNA sequencing, and Hsin-Chih Wang for technical support in molecular research. We thank the staff of the cell imaging core at the First Core Labs, National Taiwan University College of Medicine, for technical assistance. We also thank Dr. Kwang-Jen Hsiao and Dr. Chen-Jee Hong from Taipei Veterans General Hospital and Mei-Ying Liu from National Yang-Ming University for providing facilities to carry our work. We are also very grateful to the researchers and administrators of the *G. lamblia* genome database for providing genome information.

REFERENCES

- Wolfe, M. S. (1992) *Clin. Microbiol. Rev.* **5**, 93–100
- Adam, R. D. (2001) *Clin. Microbiol. Rev.* **14**, 447–475
- Ankarklev, J., Jerlström-Hultqvist, J., Ringqvist, E., Troell, K., and Svärd, S. G. (2010) *Nat. Rev. Microbiol.* **8**, 413–422
- Celiköz, A., Aciöz, M., Değerli, S., Cinar, Z., Elaldi, N., and Erandaç, M. (2005) *Pediatr. Int.* **47**, 567–571
- Gillin, F. D., Reiner, D. S., and McCaffery, J. M. (1996) *Annu. Rev. Microbiol.* **50**, 679–705
- Eichinger, D. (2001) *Curr. Opin. Microbiol.* **4**, 421–426
- Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A., and Peattie, D. A. (1989) *Science* **243**, 75–77
- Baldauf, S. L. (2003) *Science* **300**, 1703–1706
- Hashimoto, T., Nakamura, Y., Kamaishi, T., Nakamura, F., Adachi, J., Okamoto, K., and Hasegawa, M. (1995) *Mol. Biol. Evol.* **12**, 782–793
- Hirt, R. P., Logsdon, J. M., Jr., Healy, B., Dorey, M. W., Doolittle, W. F., and Embley, T. M. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 580–585
- Morrison, H. G., McArthur, A. G., Gillin, F. D., Aley, S. B., Adam, R. D., Olsen, G. J., Best, A. A., Cande, W. Z., Chen, F., Cipriano, M. J., Davids, B. J., Dawson, S. C., Elmendorf, H. G., Hehl, A. B., Holder, M. E., Huse, S. M., Kim, U. U., Lasek-Nesselquist, E., Manning, G., Nigam, A., Nixon, J. E., Palm, D., Passamaneck, N. E., Prabhu, A., Reich, C. I., Reiner, D. S., Samuelson, J., Svard, S. G., and Sogin, M. L. (2007) *Science* **317**, 1921–1926
- Best, A. A., Morrison, H. G., McArthur, A. G., Sogin, M. L., and Olsen, G. J. (2004) *Genome Res.* **14**, 1537–1547
- Seshadri, V., McArthur, A. G., Sogin, M. L., and Adam, R. D. (2003) *J. Biol. Chem.* **278**, 27804–27810
- Holberton, D. V., and Marshall, J. (1995) *Nucleic Acids Res.* **23**, 2945–2953
- Luján, H. D., Mowatt, M. R., Conrad, J. T., Bowers, B., and Nash, T. E. (1995) *J. Biol. Chem.* **270**, 29307–29313
- Mowatt, M. R., Luján, H. D., Cotten, D. B., Bowers, B., Yee, J., Nash, T. E., and Stibbs, H. H. (1995) *Mol. Microbiol.* **15**, 955–963
- Knodler, L. A., Svärd, S. G., Silberman, J. D., Davids, B. J., and Gillin, F. D. (1999) *Mol. Microbiol.* **34**, 327–340
- Sun, C. H., and Tai, J. H. (1999) *J. Biol. Chem.* **274**, 19699–19706
- Yee, J., Mowatt, M. R., Dennis, P. P., and Nash, T. E. (2000) *J. Biol. Chem.* **275**, 11432–11439
- Elmendorf, H. G., Singer, S. M., Pierce, J., Cowan, J., and Nash, T. E. (2001) *Mol. Biochem. Parasitol.* **113**, 157–169
- Sun, C. H., Palm, D., McArthur, A. G., Svärd, S. G., and Gillin, F. D. (2002) *Mol. Microbiol.* **46**, 971–984
- Davis-Hayman, S. R., Hayman, J. R., and Nash, T. E. (2003) *Int. J. Parasitol.* **33**, 1005–1012
- Sun, C. H., McCaffery, J. M., Reiner, D. S., and Gillin, F. D. (2003) *J. Biol. Chem.* **278**, 21701–21708
- Wang, C. H., Su, L. H., and Sun, C. H. (2007) *J. Biol. Chem.* **282**, 8905–8914
- Sun, C. H., Su, L. H., and Gillin, F. D. (2006) *Mol. Biochem. Parasitol.* **146**, 45–57
- Huang, Y. C., Su, L. H., Lee, G. A., Chiu, P. W., Cho, C. C., Wu, J. Y., and Sun, C. H. (2008) *J. Biol. Chem.* **283**, 31021–31029
- Pan, Y. J., Cho, C. C., Kao, Y. Y., and Sun, C. H. (2009) *J. Biol. Chem.* **284**, 17975–17988
- Chen, Y. H., Su, L. H., and Sun, C. H. (2008) *Int. J. Parasitol.* **38**, 1305–1317
- Chen, Y. H., Su, L. H., Huang, Y. C., Wang, Y. T., Kao, Y. Y., and Sun, C. H. (2008) *PLoS ONE* **3**, e3609
- Black, A. R., and Azizkhan-Clifford, J. (1999) *Gene* **237**, 281–302
- Verona, R., Moberg, K., Estes, S., Starz, M., Vernon, J. P., and Lees, J. A. (1997) *Mol. Cell. Biol.* **17**, 7268–7282
- Helin, K. (1998) *Curr. Opin. Genet. Dev.* **8**, 28–35
- Lavia, P., and Jansen-Dürr, P. (1999) *BioEssays* **21**, 221–230
- Stevens, C., and La Thangue, N. B. (2003) *Cell Cycle* **2**, 435–437
- DeGregori, J., and Johnson, D. G. (2006) *Curr. Mol. Med.* **6**, 739–748
- McClellan, K. A., and Slack, R. S. (2007) *Cell Cycle* **6**, 2917–2927
- Bagchi, S., Weinmann, R., and Raychaudhuri, P. (1991) *Cell* **65**, 1063–1072
- Bandara, L. R., and La Thangue, N. B. (1991) *Nature* **351**, 494–497
- Cobrinik, D., Whyte, P., Peeper, D. S., Jacks, T., and Weinberg, R. A. (1993)

Regulation of *cwp* Genes by *E2F1* in *Giardia lamblia*

- Genes Dev.* **7**, 2394–2404
40. Helin, K., Harlow, E., and Fattaey, A. (1993) *Mol. Cell. Biol.* **10**, 6501–6508
41. Dyson, N. (1998) *Genes Dev.* **12**, 2245–2262
42. Dimova, D. K., and Dyson, N. J. (2005) *Oncogene* **24**, 2810–2826
43. Logan, N., Delavaine, L., Graham, A., Reilly, C., Wilson, J., Brummelkamp, T. R., Hijmans, E. M., Bernards, R., and La Thangue, N. B. (2004) *Oncogene* **23**, 5138–5150
44. Attwooll, C., Lazzerini Denchi, E., and Helin, K. (2004) *EMBO J.* **23**, 4709–4716
45. Bandara, L. R., Buck, V. M., Zamanian, M., Johnston, L. H., and La Thangue, N. B. (1993) *EMBO J.* **12**, 4317–4324
46. Lees, J. A., Saito, M., Vidal, M., Valentine, M., Look, T., Harlow, E., Dyson, N., and Helin, K. (1993) *Mol. Cell. Biol.* **13**, 7813–7825
47. Buck, V., Allen, K. E., Sørensen, T., Bybee, A., Hijmans, E. M., Voorhoeve, P. M., Bernards, R., and La Thangue, N. B. (1995) *Oncogene* **11**, 31–38
48. Zhang, Y., and Chellappan, S. P. (1995) *Oncogene* **10**, 2085–2093
49. Zheng, N., Fraenkel, E., Pabo, C. O., and Pavletich, N. P. (1999) *Gene Dev.* **13**, 666–674
50. Chabouté, M. E., Clément, B., Sekine, M., Philipps, G., and Chaubet-Gigot, N. (2000) *Plant Cell* **12**, 1987–2000
51. Stevens, R., Mariconti, L., Rossignol, P., Perennes, C., Cella, R., and Bergounioux, C. (2002) *J. Biol. Chem.* **277**, 32978–32984
52. Vlieghe, K., Vuylsteke, M., Florquin, K., Rombauts, S., Maes, S., Ormenese, S., Van Hummelen, P., Van de Peer, Y., Inze, D., and De Veylder, L. (2003) *J. Cell Sci.* **116**, 4249–4259
53. Christensen, J., Cloos, P., Toftegaard, U., Klinkenberg, D., Bracken, A. P., Trinh, E., Heeran, M., Di Stefano, L., and Helin, K. (2005) *Nucleic Acids Res.* **33**, 5458–5470
54. Zalmas, L. P., Zhao, X., Graham, A. L., Fisher, R., Reilly, C., Coutts, A. S., and La Thangue, N. B. (2008) *EMBO Rep.* **9**, 252–259
55. Kosugi, S., and Ohashi, Y. (2002) *J. Biol. Chem.* **277**, 16553–16558
56. Keister, D. B. (1983) *Trans. R. Soc. Trop. Med. Hyg.* **77**, 487–488
57. Su, L. H., Lee, G. A., Huang, Y. C., Chen, Y. H., and Sun, C. H. (2007) *Mol. Biochem. Parasitol.* **156**, 124–135
58. McArthur, A. G., Morrison, H. G., Nixon, J. E., Passamaneck, N. Q., Kim, U., Hinkle, G., Crocker, M. K., Holder, M. E., Farr, R., Reich, C. I., Olsen, G. E., Aley, S. B., Adam, R. D., Gillin, F. D., and Sogin, M. L. (2000) *FEMS Microbiol. Lett.* **189**, 271–273
59. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
60. Singer, S. M., Yee, J., and Nash, T. E. (1998) *Mol. Biochem. Parasitol.* **92**, 59–69
61. Finn, R. D., Tate, J., Mistry, J., Coghill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2008) *Nucleic Acids Res.* **36**, D281–D288
62. Kaelin, W. G., Jr., Krek, W., Sellers, W. R., DeCaprio, J. A., Ajchenbaum, F., Fuchs, C. S., Chittenden, T., Li, Y., Farnham, P. J., Blanas, M. A., *et al.* (1992) *Cell* **70**, 351–364
63. Dou, Q. P., Markell, P. J., and Pardee, A. B. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3256–3260
64. Dickinson, L. A., Joh, T., Kohwi, Y., and Kohwi-Shigematsu, T. (1992) *Cell* **70**, 631–645
65. Wang, Y. T., Pan, Y. J., Cho, C. C., Lin, B. C., Su, L. H., Huang, Y. C., and Sun, C. H. (2010) *J. Biol. Chem.* **285**, 32213–32226
66. Nakai, K., and Kanehisa, M. (1992) *Genomics* **14**, 897–911
67. Ramírez-Parra, E., Xie, Q., Boniotti, M. B., and Gutierrez, C. (1999) *Nucleic Acids Res.* **27**, 3527–3533
68. Page, B. D., Guedes, S., Waring, D., and Priess, J. R. (2001) *Mol. Cell* **7**, 451–460
69. MacWilliams, H., Doquang, K., Pedrola, R., Dollman, G., Grassi, D., Peis, T., Tsang, A., and Ceccarelli, A. (2006) *Development* **133**, 1287–1297
70. Schaefer, J. B., and Breeden, L. L. (2004) *Cell* **117**, 849–850
71. Cao, L., Peng, B., Yao, L., Zhang, X., Sun, K., Yang, X., and Yu, L. (2010) *Biol. Direct.* **20**, 5–55
72. Nicolaides, N. C., Gualdi, R., Casadevall, C., Manzella, L., and Calabretta, B. (1991) *Mol. Cell. Biol.* **11**, 6166–6176
73. Robatzek, S., and Somssich, I. E. (2002) *Genes Dev.* **16**, 1139–1149
74. Frost, V., Grocott, T., Eccles, M. R., and Chantry, A. (2008) *Crit. Rev. Biochem. Mol. Biol.* **43**, 371–391
75. Johnson, D. G., Ohtani, K., and Nevins, J. R. (1994) *Genes Dev.* **8**, 1514–1525
76. Flygare, J., Benson, F., and Hellgren, D. (1996) *Biochim. Biophys. Acta* **1312**, 231–236
77. Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R. A., and Dynlacht, B. D. (2002) *Genes Dev.* **16**, 245–256
78. Karseder, J., Rotheneder, H., and Wintersberger, E. (1996) *Mol. Cell. Biol.* **16**, 1659–1667
79. Kundu, M., Guermah, M., Roeder, R. G., Amini, S., and Khalili, K. (1997) *J. Biol. Chem.* **272**, 29468–29474
80. Wang, H., Shao, N., Ding, Q. M., Cui, J., Reddy, E. S., and Rao, V. N. (1997) *Oncogene* **15**, 143–157
81. Lee, C. W., Sørensen, T. S., Shikama, N., and La Thangue, N. B. (1998) *Oncogene* **16**, 2695–2710
82. Suzuki, M., Okuyama, S., Okamoto, S., Shirasuna, K., Nakajima, T., Hachiya, T., Nojima, H., Sekiya, S., and Oda, K. (1998) *Oncogene* **17**, 853–865
83. Nip, J., Strom, D. K., Eischen, C. M., Cleveland, J. L., Zambetti, G. P., and Hiebert, S. W. (2001) *Oncogene* **20**, 910–920
84. Schlisio, S., Halperin, T., Vidal, M., and Nevins, J. R. (2002) *EMBO J.* **21**, 5775–5786
85. Zhu, W., Giangrande, P. H., and Nevins, J. R. (2004) *EMBO J.* **23**, 4615–4626
86. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* **31**, 3497–3500