Rymond, M. Rosbash, *ibid.* **50**, 593 (1987).
8. E. M. De Robertis, S. Lienhard, R. F. Parisot, *Nature* **295**, 572 (1982).
9. D. J. Forbes, T. B. Kornberg, M. W. Kirschner, *J. Cell Biol.* **97**, 62 (1983).
10. D. Tollervey and I. W. Mattaj, *EMBO J.* **6**, 469 (1987); N. Riedel, S. Wolin, C. Guthrie, *Science* **235**, 328 (1987).
11. J. Rinke, B. Appel, H. Bloker, R. Frank, R. Luhrmann, *Nucleic Acids Res.* **12**, 4111 (1984).
12. Z.-Q. Pan and C. Prives, unpublished results.
13. C. Cazenave, M. Chevrier, N. T. Thuong, C. Helene, *Nucleic Acids Res.* **15**, 10507 (1987).
14. T. J. Miller, D. L. Stephens, J. E. Mertz, *Mol. Cell. Biol.* **2**, 1581 (1982); M. P. Wickens and J. B. Gurdon, *J. Mol. Biol.* **163**, 1 (1983).
15. A. Fradin *et al.*, *Cell* **37**, 927 (1984).
16. T. Michaeli and C. Prives, *Mol. Cell. Biol.* **5**, 2019 (1985).
17. M. Zillmann, S. D. Rose, S. M. Berget, *ibid.* **7**, 2877 (1987).
18. Z.-Q. Pan and C. Prives, unpublished results.
19. J. B. Gurdon, *J. Embryol. Exp. Morphol.* **20**, 401 (1968).
20. S. Zeitlin and A. Efstratiadis, *Cell* **39**, 589 (1984).
21. E. Lund and J. E. Dahlberg, *J. Biol. Chem.* **259**, 2013 (1984); G. Westin, E. Lund, J. T. Murphy, U. Pettersson, J. E. Dahlberg, *EMBO J.* **3**, 3295 (1984).
22. M. R. Lerner and J. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5495 (1979).
23. I. W. Mattaj and E. M. De Robertis, *Cell* **40**, 111 (1985).
24. A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560 (1977).
25. We thank X. Y. Fu, J. C. S. Noble, and T. Michaeli for providing oligonucleotides, and J. E. Dahlberg for providing plasmids. We thank S. M. Mount, J. L. Manley, J. C. S. Noble, and J. J. Manfredi for their critical comments on the manuscript. This work was supported by USPHS grants CA33620 and CA46121.

11 April 1988; accepted 22 June 1988

The Yeast Cell Cycle Gene *CDC34* Encodes a Ubiquitin-Conjugating Enzyme

MARK G. GOEBL, JOHN YOCHEM,* STEFAN JENTSCH,†
JOHN P. MCGRATH, ALEXANDER VARSHAVSKY, BRECK BYERS‡

Mutants in the gene *CDC34* of the yeast *Saccharomyces cerevisiae* are defective in the transition from G_1 to the S phase of the cell cycle. This gene was cloned and shown to encode a 295-residue protein that has substantial sequence similarity to the product of the yeast *RAD6* gene. The *RAD6* gene is required for a variety of cellular functions including DNA repair and was recently shown to encode a ubiquitin-conjugating enzyme. When produced in *Escherichia coli*, the *CDC34* gene product catalyzed the covalent attachment of ubiquitin to histones H2A and H2B in vitro, demonstrating that the CDC34 protein is another distinct member of the family of ubiquitin-conjugating enzymes. The cell cycle function of *CDC34* is thus likely to be mediated by the ubiquitin-conjugating activity of its product.

THE CRUCIAL TRANSITION FROM G_1 to the S phase of the cell cycle has been the subject of detailed physiological and genetic analysis in the yeast *Saccharomyces cerevisiae*. When yeast cells reach a critical size during the G_1 phase of the cell cycle, they perform a function known as "start" and become committed to undergoing a cell division (*1, 2*). The subsequent transition to S phase entails the coordinate initiation of several events, which include bud emergence, spindle morphogenesis, and, finally, DNA replication (*1, 3*). Many of the genetically controlled functions required during these early phases of the cell cycle have been identified by characterization of temperature-sensitive (ts) mutations in the cell division cycle (CDC) genes (*1*). Several of the *cdc* mutants fail to execute the start function under nonpermissive conditions (*1, 2*). Molecular analysis of the start *CDC* genes has shown that some of their products regulate pathways of protein phosphorylation (*4*). The start function is also regulated by yeast analogs of the mammalian guanosine 5′-triphosphate (GTP)–binding proteins G_S-α (*5*) and RAS (*6*). After start, functions mediated by the *CDC4* and the *CDC34* gene products are also required before the initiation of chromosomal DNA replication can occur (*3, 7, 8*). Under nonpermissive conditions, ts mutants in *CDC4* and *CDC34* develop numerous abnormal (elongated) buds, and the spindle pole body duplicates but fails to undergo the separation required for spindle formation (*7*).

Our molecular analysis of the functions following start was initiated by characterization of the cloned *CDC4* gene. The deduced amino acid sequence of the 779-residue *CDC4* product contains a repeated motif similar to that found within the β subunits of mammalian GTP-binding proteins, such as G_S and transducin (*9*). Another segment of the *CDC4* product is similar to the products of the *CDC36* gene and the mammalian *ets* oncogene (*10*). Here we describe the molecular and functional analysis of another post-start gene, *CDC34*.

Yeast DNA complementing the ts mutation *cdc34*-1 in strain G101 was isolated from a library of *S. cerevisiae* genomic DNA (*11*) constructed in the vector YRp7 (*12*). The putative *CDC34* gene was mapped within the insert [Fig. 1, (*12*)], and found by R-loop analysis to encode a polyadenylated RNA ~1.2 kb in length (*8*). To verify that the cloned DNA sequence contained *CDC34*, we integrated the complementing plasmid into the yeast genome via homologous recombination and mapped the site of integration. Strain G102 :: *CDC34*, which contains a chromosomally integrated copy of plasmid pCDC34-79, was mated to strain G101 (*13*). Tetrads resulting from the sporulation of this diploid demonstrated tight linkage between *cdc34*-1 and the *TRP1* gene present on the integrated plasmid (19/19 parental ditype tetrads). Integration of plasmid pCDC34-79 at the *cdc34* locus indicated that the plasmid contains the *CDC34* gene.

By conventional genetic mapping (Table 1), we located *CDC34* on the right arm of chromosome IV. The map order is *CEN4-trp1-cdc34-mak21-rad55*. Since the restriction map of *CDC34* differs from that of *RRP1*, which also maps in this region, *cdc34* is not allelic to any previously mapped yeast gene.

To determine whether the distinctive multibudded morphology of arrested *cdc34* cells results from the loss of *CDC34* function under nonpermissive conditions, we inserted the yeast *HIS3* gene within the open reading frame of one copy of *CDC34* in two diploid strains (*14*). When these diploids were sporulated and dissected, 60

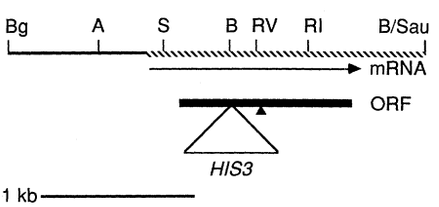

Fig. 1. Physical map of the *CDC34* locus. Indicated is part of plasmid pCDC34-79 showing positions of the sequenced yeast genomic DNA insert (striped segment), the *CDC34* mRNA (arrow), the *CDC34* open reading frame (ORF), the insertion used to generate a truncated *cdc34*-3 ORF (*19*) (small triangle), and the site of a *HIS3* insertion used to disrupt *CDC34* (*14*). Designations: A, Apa I; B, Bam HI; Bg, Bgl II; RI, Eco RI; RV, Eco RV; S, Sca I; and B/Sau, Bam HI/Sau 3 AI restriction site junction.

M. G. Goebl, J. Yochem, B. Byers, Department of Genetics, University of Washington, Seattle, WA 98195. S. Jentsch, J. P. McGrath, A. Varshavsky, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

*Present address: Department of Biology, Princeton University, Princton, NJ 08544.
†Present address: Friedrich-Miescher-Laboratorium der Max-Planck-Gesellschaft, Spemannstrasse 37-39, 7400 Tübingen, Federal Republic of Germany.
‡Send correspondence and requests for reprints to B. Byers, Department of Genetics, SK-50, University of Washington, Seattle, WA 98195.

Table 1. Chromosomal mapping of *cdc34* [see (*27*) for details].

Gene pair	Segregation (no. of tetrads)*			Map distance† (cM)
	PD	NDP	TT	
aro1-cdc34	12	1	23	42.5
trp1-rad55	31	7	108	56.9
trp1-cdc34	43	0	100	44.6
cdc34-rad55	206	0	47	9.2
cdc34-mak21	104	0	3	1.4
mak21-rad55	91	0	16	7.5

*PD, parental ditype; NPD, nonparental ditype; TT, tetratype. †Value (x') in centimorgans (cM) as described in (*28*).

tetrads each contained two viable spores (which were His$^-$) and two inviable spores. The inviable spores germinated without dividing, forming multiple, elongated buds (*8*). The striking similarity of this probable null *cdc34* phenotype to that caused by the *cdc34*-1 ts mutation indicates that the defect in *cdc34*-1 mutants results from loss of *CDC34* function under nonpermissive conditions.

When a 2.2-kb Apa I–Sau 3AI DNA fragment containing *CDC34* (Fig. 1) was hybridized to electrophoretically fractionated *S. cerevisiae* DNA that had been digested with various restriction endonucleases, several cross-hybridizing bands were detected (Fig. 2A). A computer-assisted search of the sequenced *CDC34* DNA (see below, Fig. 3) revealed no sequence similarities between the probe and known yeast repetitive elements. However, the DNA encoding a highly acidic region of the *CDC34* product consists largely of aspartic acid–coding triplets GAC and GAT in tandem array, and could possibly cause cross-hybridization to other genes encoding tracts of aspartic acid residues. To explore this possibility, the Apa I–Sau 3AI fragment (Fig. 1) was digested with Eco RV to yield two fragments, one of which contained all of the repeated $(GAC/T)_n$ sequences (Fig. 3). Both of these fragments hybridized to *CDC34* sequences (Fig. 2B), whereas only the fragment containing the $(GAC/T)_n$ repeats also hybridized to multiple DNA sequences (*8*). We conclude that although a number of yeast genes may encode tracts rich in aspartic acid, there are no close homologs of *CDC34* in the *S. cerevisiae* genome.

The nucleotide sequence of a 1820-bp genomic DNA segment from pCDC34-79 that includes *CDC34* is shown in Fig. 3. The only large open reading frame encodes a protein of 295 residues if the first methionine codon of the frame serves as the initiation codon. We conclude that this open reading frame encodes the *CDC34* gene product, which would have a molecular mass of 34.1 kD. A computer-assisted search (*15*) for similarities between the deduced amino acid sequence of the *CDC34* product and known proteins revealed substantial similarity only to the yeast RAD6 protein (Fig. 3). The *RAD6* gene is required for a variety of functions including DNA repair, induced mutagenesis, and sporulation (*16*). Amino acid sequence similarities between the *RAD6* and *CDC34* products are distributed throughout their lengths, and are especially pronounced in the central region of the *CDC34* product as well as in the COOH-terminal regions where aspartic acid–rich tracts occur in both proteins. Overall, 38% of the amino acids in the *RAD6* product are identical to those in the *CDC34* product. With the inclusion of conservative replacements (*17*), 55% of the *RAD6* product is conserved within the *CDC34* product. Recently, Jentsch *et al.* (*18*) found that *RAD6* is a member of a family of genes encoding related but distinct ubiquitin-conjugating enzymes (E2 enzymes) that catalyze the covalent attachment of ubiquitin, a highly conserved 76-residue protein, to specific protein substrates (*18*). The striking amino acid sequence similarity between the *RAD6* and *CDC34* gene products (Fig. 3) suggested that the latter protein may also be a ubiquitin-conjugating enzyme.

To test this possibility, we carried out assays for ubiquitin-conjugating activity in extracts of *E. coli* expressing *CDC34*, as well as in control extracts of *E. coli* harboring the same vector but with *CDC34* inserted in the opposite (presumably non-expressing) orientation (*19, 20*). The conjugation of ubiquitin to acceptor proteins begins with an adenosine 5′-triphosphate (ATP)–requiring step in which the COOH-terminal glycine residue of ubiquitin is joined, through a high-energy thiolester bond, to a cysteine residue of the ubiquitin-activating enzyme, E1 (*21–23*). The activated ubiquitin is then transferred to a specific cysteine residue within a ubiquitin-conjugating (E2) enzyme, which in turn catalyzes the formation of a branched isopeptide bond–mediated conjugate between the COOH terminus of ubiquitin and the ε-amino group of a lysine residue in the acceptor protein (*21, 23*). When supplemented with ATP and purified yeast ubiquitin-activating enzyme, E1, only extracts of *CDC34*-expressing *E. coli* cells mediated the covalent conjugation of ^{125}I-labeled ubiquitin to histones H2B (Fig. 4A) and H2A (*24*). The formation of a metastable intermediate, in which an E2 enzyme is joined to ubiquitin by a thiolester bond in a reaction dependent on both ATP and ubiquitin-activating (E1) enzyme, could also be demonstrated with the extract of *E. coli* expressing the *CDC34* gene but not with an otherwise identical extract lacking the CDC34 protein (Fig. 4, B and C). The remarkable degree of substrate specificity of the CDC34 ubiquitin-conjugating enzyme is indicated by its apparent failure to ubiquitinate endogenous proteins in the *E. coli* extract (Fig. 4A).

The demonstration of ubiquitin-histone conjugating activity in bacterial cells expressing *CDC34* and the ability of the CDC34 protein to form a thiolester bond–mediated covalent complex with ubiquitin directly identifies the *CDC34* gene as yet another member of the gene family encoding related but distinct E2 enzymes. In

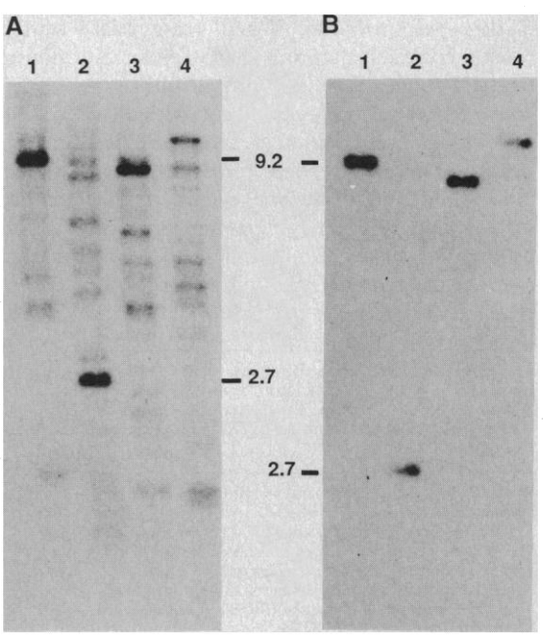

Fig. 2. DNA hybridization analysis of yeast genomic DNA with *CDC34* probes. *Saccharomyces cerevisiae* DNA was digested with the following restriction endonucleases before electrophoresis in a 1.0% agarose gel and DNA hybridization with ^{32}P-labeled probes (*29, 30*): lanes 1, Bam HI; lanes 2, Bgl II; lanes 3, Cla I; and lanes 4, Eco RI. (**A**) The hybridization probe was a 2.2-kb Apa I–Sau 3AI fragment (see Fig. 1) containing the entire *CDC34* gene. (**B**) The hybridization probe was a 1.0-kb Apa I–Eco RV fragment containing the portion of *CDC34* that encodes an NH_2-terminal part of the CDC34 protein (see Figs. 1 and 3).

previous work, a yeast gene encoding the ~30-kD $E2_{30K}$ enzyme was designated *UBC1* [for *ub*iquitin-*c*onjugating enzyme, (*18*)]. The *RAD6* gene, which encodes the $E2_{20K}$ enzyme, was renamed *UBC2* (*18*). We now propose to rename *CDC34*, which encodes the $E2_{34K}$ enzyme, as *UBC3*.

Earlier studies with enzymes from mammalian reticulocytes have indicated that ubiquitination by some of the E2 enzymes requires another distinct protein, E3 (*21*). E3-dependent ubiquitin conjugation by an E2 enzyme appears to involve specifically those protein substrates that are targeted for degradation through the ubiquitin-dependent proteolytic pathway (*21, 22*). An alternative role for ubiquitin, suggested by the relative metabolic stability of ubiquitinated derivatives of actin and integral membrane proteins (*25*), as well as histones (*26*), is that the reversible joining of ubiquitin to an acceptor protein could modulate protein function without metabolically destabilizing the acceptor protein (*21*). The ubiquitin-histone conjugation by either the RAD6 (UBC2) or CDC34 (UBC3) enzymes is E3-independent (Fig. 4) (*18*), consistent with the possibility that these enzymes are not a part of the ubiquitin-dependent proteolytic pathway (*21*).

In addition to their similar substrate requirements in vitro, both the RAD6 (UBC2) and CDC34 (UBC3) enzymes have strikingly acidic COOH-terminal regions (Fig. 3). These regions, which are absent from the other sequenced ubiquitin-conjugating enzyme, *UBC1* (*24*), might be involved in ionic interactions with basic protein substrates such as histones. To address this possibility, we introduced a stop codon into the *CDC34* (*UBC3*) reading frame immediately downstream of the region where the amino acid sequences are similar in all three sequenced E2 enzymes, UBC1 to UBC3 [(*19, 24*); see also Fig. 3]. When produced in *E. coli*, the corresponding truncated cdc34 protein (encoded by *cdc34*-3), which lacks the entire acidic COOH-terminal region (Fig. 3), was able to form a thiolester bond–mediated covalent complex with ^{125}I-labeled ubiquitin in a reaction dependent on the presence of both ATP and purified ubiquitin-activating enzyme E1 (Fig. 4, B and C, lanes 4). This complex migrates in a polyacrylamide-SDS gel at ~28 kD (Fig. 4, B and C, lanes 4; compare with lanes 2), consistent with the addition of an ~8-kD ubiquitin moiety to the ~20-kD truncated cdc34 (ubc3) protein. However, when assayed for ubiquitin-histone conjugation, the truncated protein, unlike the wild-type CDC34 (UBC3) enzyme, exhibits no detectable ubiquitin-histone conjugating activity (Fig. 4A, lane 4). These results strongly suggest that at least the in vitro substrate specificity of the CDC34 (UBC3) enzyme is due in part to ionic interactions between its acidic COOH-terminal region and basic protein substrates.

The enzymatic activity of the CDC34 (UBC3) protein suggests that its cell cycle function is mediated by ubiquitination of specific target proteins. What might these targets be? Our experiments show that histones H2A and H2B, but not endogenous proteins in an *E. coli* extract, can serve as in vitro substrates for ubiquitination by the CDC34 (UBC3) enzyme (Fig. 4A). In addition, the highly acidic stretch at the COOH terminus of the CDC34 (UBC3) is involved in the recognition of histone substrates in vitro by this E2 enzyme but is not required for the formation of ubiquitin-E2 thiolester intermediates. These results suggest that the in vivo substrates of the CDC34 (UBC3) enzyme may be few in number and may include DNA-bound histones. Furthermore, since the CDC34 (UBC3) protein is an E3-independent ubiquitin-conjugating enzyme, at least with the histones as in vitro substrates (see above), its in vivo substrates are unlikely to be metabolically destabilized by CDC34 (UBC3)–mediated ubiquitination. The CDC34 enzyme may instead modulate substrate function either directly or via interactions with other proteins that distinguish between ubiquitinated and nonubiquitinated substrate. The phenotypic similarities shared by *cdc34* (*ubc3*) and *cdc4* mutants suggest that the wild-type product of *CDC4* (*9*) participates with the

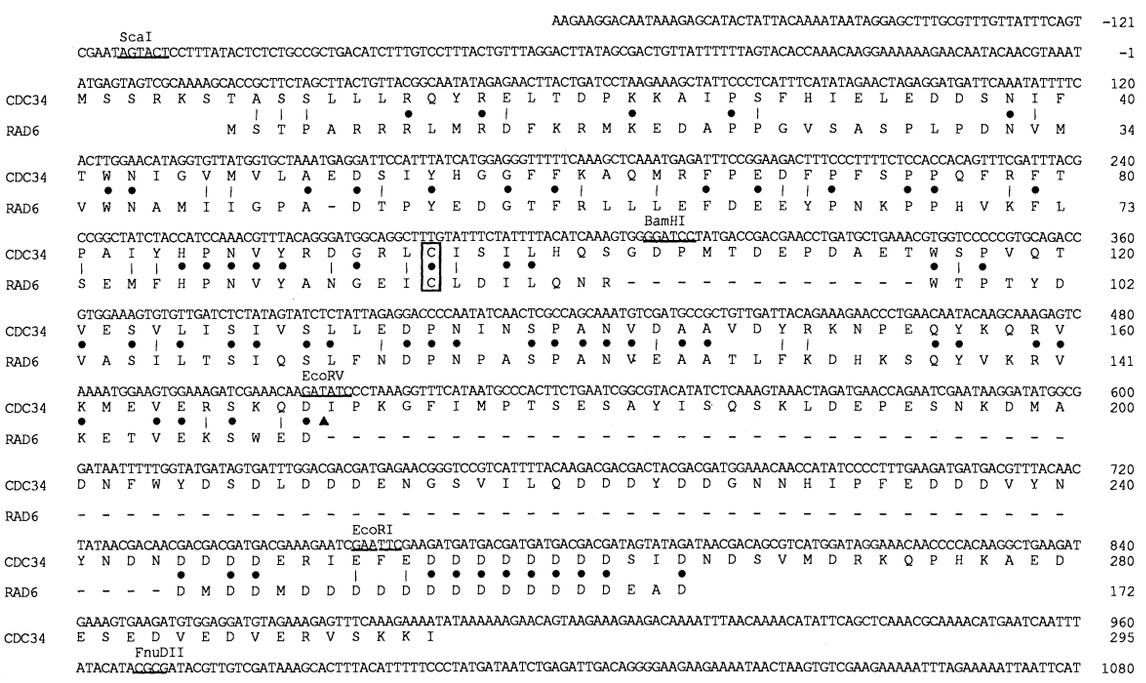

Fig. 3. Nucleotide sequence of the *S. cerevisiae* *CDC34* gene and the deduced amino acid sequence of its product compared to that of the *RAD6* product. Some restriction endonuclease cleavage sites are underlined. Numbering of the *CDC34* nucleotide sequence (*15*) begins at the A residue of the first methionine codon in the open reading frame, and numbering of deduced amino acid sequence of the *CDC34* product begins at the same methionine. The alignment shown maximizes the frequency of sequence identities with the RAD6 protein [(*16*), see text]. Identical residues in the two proteins are indicated by dots; vertical lines indicate conservative amino acid replacements grouped according to Dayhoff (*17*) as follows: C; S,T,P,A,G; N,D,E,Q; H,R,K; M,I,L,V; F,W,Y. The single cysteine residues of the CDC34 and RAD6 proteins that are expected to be essential for the ubiquitin-conjugating activity of these enzymes (*21–23*) are boxed. An arrowhead denotes the site at which a stop codon was inserted into the wild-type reading frame to generate the truncated protein encoded by *cdc34*-3 (*19*). Single-letter amino acid abbreviations: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.

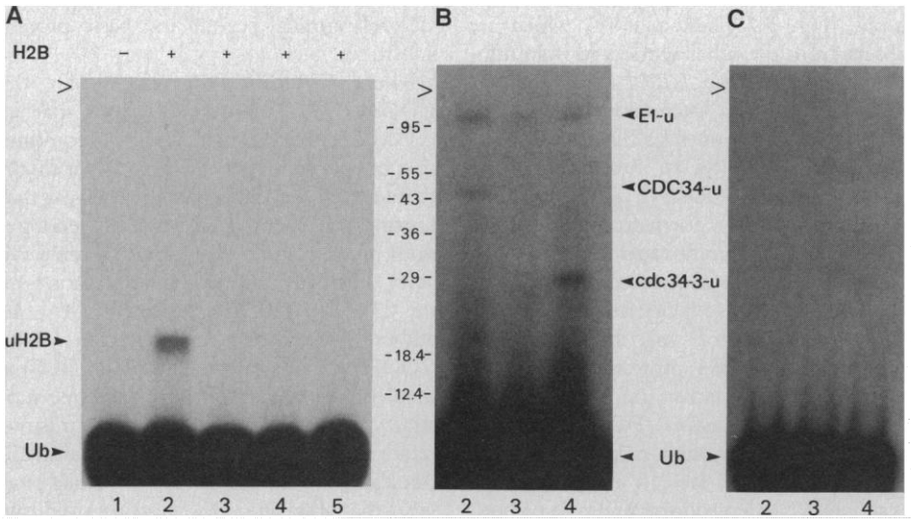

Fig. 4. The *CDC34* gene product is a ubiquitin-conjugating enzyme. Extracts of induced *E. coli* cells expressing the *CDC34* gene (lanes 1, 2, and 5), extracts of cells containing the *CDC34* gene in an opposite (non-expressing) orientation within the vector (lanes 3), and extracts of cells expressing the mutant (truncated) *cdc34-3* gene (lanes 4) were assayed for ubiquitin conjugation activity (*20*). Products of the assay were separated by electrophoresis in an 18% polyacrylamide-SDS gel and detected by autoradiography (*20*). (**A**) Extracts were assayed for ubiquitin-histone H2B conjugation in either the absence (lane 5) or presence (lanes 1 to 4) of the purified ubiquitin-activating enzyme, E1, and in the presence of ATP, ^{125}I-labeled ubiquitin, and histone H2B (*20*). The extract of *CDC34*-expressing cells could form a monoubiquitinated species of histone H2B in an E1-dependent reaction (lanes 2 and 5), whereas neither the extract of non-expressing cells (lane 3) nor the extract of cells expressing the mutant (truncated) *cdc34-3* gene (lane 4) possessed this activity. (**B**) When assayed for ubiquitin–CDC34 protein thiolester formation (*21–23*), complexes with the properties of thiolesters were detected between ubiquitin and added E1 (E1~u; lanes 2 to 4), CDC34 protein (CDC34~u, ~45 kD; lane 2), and cdc34-3 protein (cdc34-3~u, ~28 kD; lane 4). (**C**) Same as in (B) but the samples were heated under reducing conditions (*20*) before electrophoresis to selectively cleave thiolester-mediated complexes. Designations: Ub, ubiquitin; uH2B, monoubiquitinated histone H2B. Molecular masses of size markers are in kilodaltons. Arrowheads indicate gel origins.

CDC34 (UBC3) ubiquitin-conjugating enzyme in mediating the G_1 to S transition.

Although the *RAD6* (*UBC2*) and *CDC34* (*UBC3*) products are clearly distinct both structurally and functionally, their enzymatic specificities appear to be quite similar in vitro: both enzymes apparently fail to ubiquinate endogenous proteins in an *E. coli* extract but efficiently ubiquitinate histones H2A and H2B in E3-independent reactions (Fig. 4) (*18*). If the CDC34 (UBC3) and RAD6 (UBC2) enzymes share functionally relevant in vivo substrates, the distinct functions of these enzymes must be derived from differences in their specific modes of action. One possibility is that ubiquitination of histones or other substrates by the CDC34 (UBC3) and RAD6 (UBC2) enzymes might be confined to different specific regions of chromatin, such as sites involved in the initiation of DNA replication in the case of the CDC34 (UBC3) enzyme or sites of DNA damage in the case of the RAD6 (UBC2) enzyme. Alternatively, the functionally relevant in vivo protein substrates of the CDC34 (UBC3) and RAD6 (UBC2) enzymes may include factors involved in the regulation of genes that control, respectively, post-start events required for the initiation of DNA replication and events necessary for DNA repair or sporulation. All of the above possibilities are consistent with immunofluorescence evidence (*8*) that the CDC34 (UBC3) enzyme is a nuclear protein.

That progress through the cell cycle may involve specific ubiquitination events has been suggested by earlier work on the mutant mammalian cell line, ts85, whose preferential G_2 arrest under nonpermissive conditions results from thermolability of the ubiquitin-activating enzyme in these cells (*22*). The discovery that a regulatory step in the G_1 to S transition is controlled in yeast by a specific ubiquitin-conjugating enzyme underscores the remarkable functional diversity of the ubiquitin system and opens up a new direction for studies of the cell cycle.

REFERENCES AND NOTES

1. L. H. Hartwell, J. Culotti, J. R. Pringle, B. J. Reid, *Science* **183**, 46 (1974); J. R. Pringle and L. H. Hartwell, in *The Molecular Biology of the Yeast* Saccharomyces: *Life Cycle and Inheritance*, J. N. Strathern, E. W. Jones, J. E. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981), pp. 97–142; D. Botstein and R. Maurer, *Annu. Rev. Genet.* **16**, 61 (1982).
2. S. I. Reed, *Genetics* **95**, 561 (1980).
3. L. H. Hartwell, *J. Bacteriol.* **115**, 966 (1973); L. M. Hereford and L. H. Hartwell, *J. Mol. Biol.* **84**, 445 (1974); J. Culotti, thesis, University of Washington, Seattle (1974).
4. S. I. Reed, J. A. Hadwiger, A. T. Lorincz, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4055 (1985); J. Lisziewicz, A. Godany, H. H. Forster, H. Kunzel, *J. Biol. Chem.* **262**, 2549 (1987).
5. C. Dietzel and J. Kurjan, *Cell* **50**, 1001 (1987); I. Miyajima *et al.*, *ibid.*, p. 1011.
6. T. Toda *et al.*, *ibid.* **40**, 27 (1985).
7. B. Byers and L. Goetsch, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 123 (1973); B. Byers, in *Molecular Genetics in Yeast: Alfred Benzon Symposium 16*, D. von Wettstein, J. Friis, M. Kielland-Brandt, A. Stenderup, Eds. (Munksgaard, Copenhagen, 1981), pp. 119–133.
8. B. Byers, M. Goebl, L. Goetsch, J. Yochem, unpublished data.
9. H. K. W. Fong *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2162 (1986); J. Yochem and B. Byers, *J. Mol. Biol.* **195**, 233 (1987).
10. T. A. Peterson *et al.*, *Nature* **309**, 556 (1984).
11. K. A. Nasmyth and S. I. Reed, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2119 (1980).
12. The vector YRp7, which consists of the *S. cerevisiae TRP1-ARS1* sequence inserted into the Eco RI site of pBR322, is able to replicate in both *S. cerevisiae* and *E. coli* [K. Struhl *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1035 (1979)]. All plasmids were constructed by standard techniques (*31*). The plasmids were maintained in the *E. coli* strains DH-1 or RR1. The yeast strains used were G101 (*MATα cdc34-1 trp1 ura1 ade2 his*) and G102 (*MATα cdc34-1 trp1 ura1 leu2 lys2*). Plasmid DNA was isolated from *E. coli* by alkaline lysis [H. C. Birnboim and J. Doly, *Nucleic Acids Res.* **7**, 1513 (1979)] and from *S. cerevisiae* as described (*11*). Transformations of *E. coli* and *S. cerevisiae* were performed by standard techniques [M. Mandel and A. Higa, *J. Mol. Biol.* **53**, 159 (1970); and (*32*)]. Although the *cdc34-1* mutant is viable at 37°C on medium containing 1*M* sorbitol, it grows more slowly than *CDC34* strains. Thus, *cdc34-1* cells transformed with *CDC34* could be initially identified by their ability to grow more rapidly under the above conditions (8).
13. The plasmid pCDC34-79 was transformed into the strain G102 (*32*), and a stable ts$^+$ (non-ts) transformant was isolated and designated G102 :: *CDC34* [after confirmation by DNA hybridization analysis (*29, 30*) that the plasmid integration site was within the homologous genomic sequences]. Standard genetic techniques were used to construct and sporulate diploids for tetrad analysis (*32*).
14. To construct a disruption of *CDC34* [R. Rothstein, *Methods in Enzymol.* **101**, 202 (1983)], the plasmid CMp170 was digested with Bam HI, liberating a 1.7-kb fragment containing the yeast *HIS3* gene [K. Struhl, *Nucleic Acids Res.* **13**, 8587 (1985)]. This fragment was inserted, by standard techniques (*31*), into the Bam HI site of pGEM34 H/S (*33*). The resulting plasmid was digested with Apa I–Eco RI (Fig. 1) and transformed (*32*) into two diploid strains. DNA hybridization analysis of DNA from His$^+$ transformants and from the spore-derived colonies of one tetrad from each transformant confirmed that His$^+$ diploids contained both the wild-type and insertional alleles of *CDC34*, whereas the viable spore colonies contained only the wild-type allele of *CDC34*.
15. The nucleotide sequence of *CDC34* was determined by the chain termination method [F. Sanger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977)]. The plasmids pGEM34 B/S and pGEM34 H/S (*33*) were used to generate a series of overlapping deletions extending into either side of *CDC34* by recloning fragments produced by controlled exonuclease III digestion [S. Henikoff, *Gene* **28**, 351 (1984)]. This approach allowed both strands to be sequenced completely. The similarity between the deduced amino acid sequences of CDC34 and RAD6 proteins was found by searching a database of amino acid sequences of yeast proteins (collected by M. Goebl) by means of the GENEPRO program (Riverside Scientific Enterprises, Seattle). Alignment of the two gene products utilized the algorithm of W. J. Wilbur and D. Lipman [*Proc. Natl. Acad. Sci. U.S.A.* **80**, 726 (1983)].
16. P. Reynolds, S. Weber, L. Prakash, *Proc. Natl. Acad.*

Sci. U.S.A. **82**, 168 (1986); R. H. Haynes and B. A. Kunz, in *The Molecular Biology of the Yeast* Saccharomyces: *Life Cycle and Inheritance*, J. N. Strathern, E. W. Jones, J. E. Broach, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981), pp. 371–414; C. W. Lawrence, *Adv. Genet.* **21**, 173 (1982).
17. M. O. Dayhoff, in *Atlas of Protein Sequence and Structure*, M. O. Dayhoff, Ed. (National Biomedical Research Foundation, Washington, DC, 1978), vol. 5, suppl. 3, p. 345.
18. S. Jentsch, J. P. McGrath, A. Varshavsky, *Nature* **329**, 131 (1987).
19. The *CDC34*-containing Sca I–Fnu DII fragment (Fig. 3) was subcloned with Hind III linkers (New England Biolabs) into the Hind III site of the vector pKK223-3 [J. Brosius and J. R. Lupski, *Methods Enzymol.* **153**, 54 (1987)], downstream of the vector's *tac* promoter, in both the expressing and opposite (presumably nonexpressing) orientations to generate plasmids pK34-1 and pK34-2, respectively. To construct pK34-3, a single copy of a Bgl II linker, CAGATCTG (New England Biolabs), was inserted into plasmid pK34-1 at its unique Eco RV site (see Fig. 3), yielding the mutation *cdc34*-3, which encodes a truncated protein. Insertion of the linker generates a Gln-Ile insertion at amino acid residue 170 of the wild-type CDC34 protein (Fig. 3), followed (at the nucleotide sequence level) by a TGA stop codon.
20. *Escherichia coli* cells harboring the pK34-1, pK34-2, or pK34-3 plasmid (*19*) were induced (*19*) and lysed by freeze-thawing in 5 volumes of 50 m*M* tris-HCl (*p*H 7.5), 1 m*M* EDTA, and 1 m*M* dithiothreitol (DTT). The assay for ubiquitin-histone conjugation was carried out in 50 μl of 50 m*M* tris-HCl (*p*H 7.5), 2 m*M* ATP, 5 m*M* $MgCl_2$, and 0.2 m*M* DTT, in the presence of 50 pmol of ^{125}I-labeled ubiquitin ($\sim 1 \times 10^5$ cpm) (*21, 22*), 0.2 μg of bovine histone H2B (Boehringer Mannheim), and 4 pmol of the purified yeast E1 enzyme (*18*). The assay for ubiquitin thiolesters was carried out as above except that histone H2B was omitted, and the reaction products were incubated for 5 min at room temperature in an SDS-containing electrophoretic sample buffer lacking reducing agents before a polyacrylamide–SDS gel electrophoresis at 4°C (*22*). Reduction of thiolesters was performed by heating the samples prior to electrophoresis at 100°C for 10 min in an SDS sample buffer containing 4*M* 2-mercaptoethanol (*18, 21*).
21. D. Finley and A. Varshavsky, *Trends Biochem. Sci.* **10**, 343 (1985); A. Hershko and A. Ciechanover, *Prog. Nucleic Acid Res. Mol. Biol.* **33**, 19 (1986); M. Rechsteiner, *Annu. Rev. Cell Biol.* **3**, 1 (1987).
22. D. Finley, A. Ciechanover, A. Varshavsky, *Cell* **37**, 43 (1984); A. Ciechanover, D. Finley, A. Varshavsky, *ibid.*, p. 57; A. Bachmair, D. Finley, A. Varshavsky, *Science* **234**, 179 (1986); D. Finley, E. Özkaynak, A. Varshavsky, *Cell* **48**, 1035 (1987).
23. C. M. Pickart and I. A. Rose, *J. Biol. Chem.* **261**, 10210 (1985).
24. S. Jentsch, J. P. McGrath, A. Varshavsky, unpublished data.
25. E. Ball *et al.*, *Cell* **51**, 221 (1987); T. Yarden *et al.*, *Nature* **323**, 226 (1986); M. Siegelman *et al.*, *Science* **231**, 823 (1986).
26. H. Busch, *Methods Enzymol.* **106**, 238 (1984); R. D. Mueller, H. Yasuda, C. L. Hatch, W. M. Bonner, E. M. Bradbury, *J. Biol. Chem.* **260**, 5147 (1985); L. Levinger and A. Varshavsky, *Cell* **28**, 375 (1982).
27. The chromosome bearing the *CDC34* gene was identified by probing a nitrocellulose filter containing chromosome-sized yeast genomic DNA separated by pulsed-field electrophoresis [G. Carle and M. V. Olson, *Nucleic Acids Res.* **12**, 5647 (1984)] with ^{32}P-labeled pCDC34-79 DNA (*29*). The results (*8*) tentatively localized *CDC34* to chromosome IV. Diploid strains containing *cdc34*-1 and various genetic markers present on the right arm of chromosome IV (*32*) were then constructed. The results of tetrad analyses with these strains are presented in Table 1.
28. R. K. Mortimer and D. Schild, *Microbiol. Rev.* **49**, 181 (1985).
29. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977). Hybridization conditions [G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984)]: overnight incubation in 7% SDS, 1 m*M* EDTA, and 0.5*M* sodium phosphate (*p*H 7.2) at 65°C, followed by a 5-min washing with 5% SDS and 40 m*M* sodium phosphate (*p*H 7.2) at 65°C, and further 5-min washings (five times) with 1% SDS and 40 m*M* sodium phosphate (*p*H 7.2) at room temperature.
30. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
31. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
32. F. Sherman, G. R. Fink, J. Hicks, *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1981).
33. A 3.2-kb Bgl II–Sph I fragment and a 3.3-kb Hind III–Sal I fragment from pCDC34-79 (Fig. 1) were subcloned into pGEM-1 (Promega Biotech, Madison, WI); the resulting plasmids were designated pGEM34 B/S and pGEM34 H/S, respectively.
34. We thank L. Hartwell, W. Fangman, A. Hopper, J. Hurley, D. Finley, and B. Bartel for advice during the preparation of the manuscript. We thank C. Mann for the gift of plasmid CMp170 and R. McCarroll for the Southern blot of separated yeast chromosomes. Supported by grants from the National Institute of General Medical Sciences to B.B. (GM18541) and A.V. (GM31530).

9 March 1988; accepted 29 June 1988

Physical Analysis of Transcription Preinitiation Complex Assembly on a Class II Gene Promoter

MICHAEL W. VAN DYKE, ROBERT G. ROEDER, MICHÈLE SAWADOGO

Transcription of protein-encoding genes by human RNA polymerase II requires multiple ancillary proteins (transcription factors). Interactions between these proteins and the promoter DNA of a viral class II gene (the major late transcription unit of adenovirus) were investigated by enzymatic and chemical footprinting. The experiments indicated that the assembly of functionally active RNA polymerase II–containing transcription preinitiation complexes requires a complete set of transcription factors, and that both specific protein-DNA and protein-protein interactions are involved. This allows individual steps along the transcription reaction pathway to be tested directly, thus providing a basis for understanding basic transcription initiation mechanisms as well as the regulatory processes that act on them.

GENE ACTIVITY IN EUKARYOTES IS frequently regulated at the level of DNA transcription. Understanding the molecular mechanism behind this transcriptional control of gene expression has become one of the major goals of modern biology. Mutational analysis has revealed that the activity of protein-encoding genes is governed by several kinds of promoter elements. A basal level of transcription is usually observed when all but a small region of the promoter DNA has been deleted. This minimum or core promoter is often centered on a TATA box sequence (*1*). Two classes of cis-acting DNA elements are involved in modulating the TATA box–driven transcription. Upstream promoter elements act at short distance, while enhancer elements can be located up to several thousand base pairs from the transcription initiation site (*2*). The various components of the human class II transcription machinery are being elucidated by the biochemical dissection of crude cell-free systems (*3, 4*). This approach has revealed that specific transcription initiation by RNA polymerase II requires the coordinated action of multiple protein factors. These can be classified into two categories: general transcription factors and upstream element-binding proteins (*2*). Through both kinetic analysis and the use of various inhibitors, several steps have been defined within the transcription mechanism (*5–7*). These include: (i) a commitment of the template after binding of a subset of the transcription factors, (ii) formation of an activated state through the action of other transcription factors and RNA polymerase II, (iii) fulfillment of an energy requirement (hydrolysis of the β-γ phosphate bond of either ATP or dATP), (iv) initiation of transcription (formation of the first phosphodiester bond), and (v) transcription elongation, with some transcription components remaining committed to the template. Within these various steps, the exact roles of the different protein factors, their assembly into active transcription complexes, and the various transitions that lead to the initiation event, remain for the most part unknown.

The major late (ML) promoter of adenovirus provides a useful model system for in vitro analysis of specific transcription by human RNA polymerase II. This promoter is quite simple structurally, with only two essential DNA elements: a TATA box at position −28 and a single upstream element at position −58 (*8–11*). At least three transcription factors (designated TFIIB, TFIID,

Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021.