Volume 15 Number 20 1987 Nucleic Acids Research

Metal-binding, nucleic acid-binding finger sequences in the *CDC16* gene of *Saccharomyces cerevisiae*

Tateo Icho and Reed B.Wickner

Section on Genetics of Simple Eukaryotes, Laboratory of Biochemical Pharmacology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 8, Room 209, Bethesda, MD 20892, USA

Received September 2, 1987; Accepted September 22, 1987

ABSTRACT

The CDC16 gene is involved in the process of chromosome segregation in mitosis and a cdc16^{ts} mutant accumulates the predominant microtubule-associated protein at the nonpermissive temperature. We find that the CDC16 gene open reading frame (ORF) is capable of encoding a protein whose calculated molecular weight and pI are 94,967 and 6.60, respectively. This hypothetical protein contains 16 cysteine residues; five are clustered at the N-terminal, 4 are placed about 3 residues apart in the middle of the peptide, and 3 are located close to the C-terminal. Each of these could form a metal-binding, nucleic acid-binding domain, suggesting this protein acts either as a repressor of the microtubule-associated protein gene or as a component necessary for spindle elongation, possibly interacting with the DNA. The start of the CDC16 ORF is only 95 bp downstream from the end of the MAK11 ORF. In this region there are two TATA boxes in tandem, but there is no room for a UAS or other regulatory sequences. An ATG is present 5 bp upstream of the start of the large ORF. Its frame terminates after only two amino acids.

INTRODUCTION

One of the most interesting, but still not well understood, processes in the yeast mitotic cell division cycle is that by which duplicated chromosomes are pulled to the poles of the spindle by the spindle fibers, thus delivering one copy of each chromosome to each of the daughter cells. Following emergence of a new bud and DNA replication, the nucleus migrates into the neck region between the mother and daughter cell. Mitosis and nuclear division then ensue as the spindle elongates with one pole in each progeny cell, the chromosomes are partitioned, and the two daughter nuclei form (1).

The cdc16-1 mutant (CDC = Cell Division Cycle) arrests at the beginning of the actual mitosis process. The cells have a single bud and the nuclei have migrated to the neck between the mother cell and daughter cell, but the spindle has not elongated and the chromosomes have not been partitioned (2, 3). The cdc16-1 temperature-sensitive mutation results in an elevated frequency of chromosome loss when mutant cells are grown at the maximum permis-

sive temperature, but no increase in mitotic recombination frequency is seen under the same conditions (4).

All the above results point to a role for the CDC16 product in the structure, formation, or functioning of the mitotic spindle. We have cloned the CDC16 gene (T. Icho and R. B. Wickner, submitted) and report here its DNA sequence analysis. The striking features of the CDC16 gene are the presence of three typical metal-binding, nucleic acid-binding finger sequences, such as first defined for TFIIIA (5), a run of threonine residues, the virtual absence of space for an upstream control region, and an ATG 5 bp upstream of the main CDC16 ORF which could limit the level of translation of CDC16 mRNA.

MATERIALS AND METHODS

Strains, Media, and Plasmids

Yeast strain YTI32 (MATa ura3 ade1 mak11-1 cdc16-1 L-A-HN M-o) was used to test the CDC16-complementing ability of various plasmids. The diploid strain YTI36 x YTI81 (MATα ura3 mak11-1 cdc16-1 ade1 his4 x MATa his6 ura3) was the parent strain for the gene disruption experiments. Escherichia coli HB101 was used for the propagation of plasmids, and JM109 was used as the host for M13. Subclones were introduced into yeast by the lithium acetate transformation method (6) using pTI15, a CEN vector with a multiple cloning site (T. Icho and R. B. Wickner, submitted).

DNA Sequencing

Plasmid deletions made by Bal31 (7) were cloned into M13mp18 or M13mp19 (8) as diagrammed in Fig. 1. DNA sequencing was carried out using the dideoxy method with [^{35}S]dATP (9, 10).

RESULTS

Deletions of the CDC16 Gene

A CDC16 plasmid, pTIC21, was subcloned from a plasmid pTIC1, which contains both the MAK11 and CDC16 genes (T. Icho and R. B. Wickner, submitted). From pTIC21, deletions were created by nuclease Bal31 and cloned into M13 (Fig. 1). DNA sequencing of 57 M13 clones, reading 400 bases per clone on the average, were compiled for the sequence. Some critical deletion fragments in M13 were recloned into a multiple cloning site CEN plasmid, pTI15 (T. Icho and R. B. Wickner, submitted), using the common SstI-SalI sites, and their complementing activity was tested by transforming the plasmids into a cdc16-1 tester strain. The growth of transformants at 37°C indicated positive complementation. These deletions defined the start and the end of the CDC16 gene as summarized in Fig. 2.

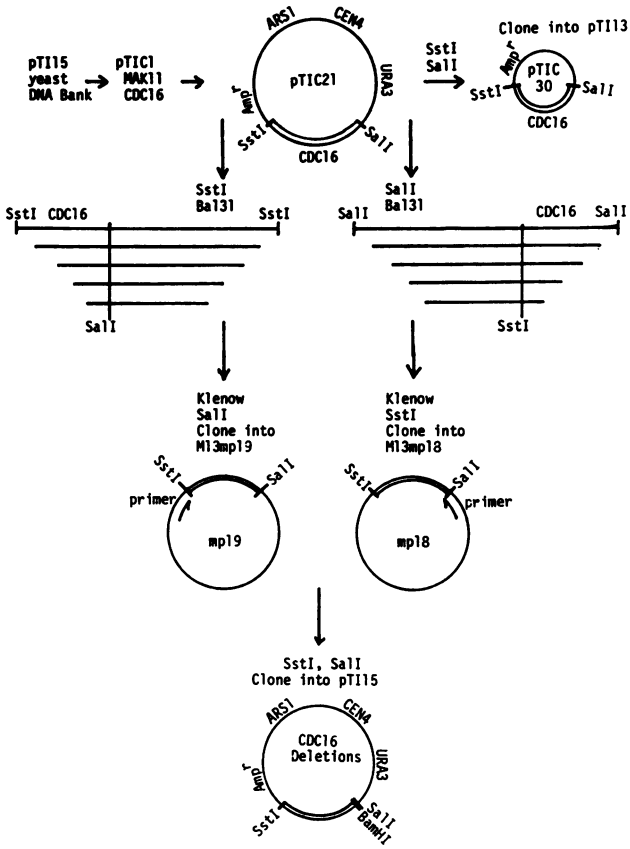


FIG. 1. The CDC16 plasmid, pTIC21, was digested with SalI, and a series of deletions were created by nuclease Bal31, repaired by the Klenow fragment of DNA polymerase I, digested with SstI, and inserted into SstI-SmaI-digested M13mp18RF. For the opposite orientation, the same plasmid was digested with SstI, and deletions were inserted between the SalI and the SmaI sites of M13mp19RF. The critical deletions in M13 were recloned into a multiple cloning site CEN vector, pTI15, using the common SstI-SalI sites, and the complementation activity for cdc16-1 was checked as summarized in Fig. 2. The plasmid pTIC30, which was used for the CDC16 gene disruption, was constructed by cloning the SstI-SalI fragment into pTI13 (26), a derivative of pAT135 carrying the same multiple cloning sites as pTI15.

Sequence of the CDC16 Gene

The DNA sequence of the CDC16 gene in plasmid pTIC21 is shown in Fig. 3. The total sequence is 3412 bp. The sequence of the first 813 bp PvuII fragment is identical with the sequence of the same region in the MAK11 plasmid, pTIC19 (T. Icho and R. B. Wickner, submitted). Starting from the 233rd base, there is a large ORF of 2523 bp which corresponds to a protein of 840 amino acid

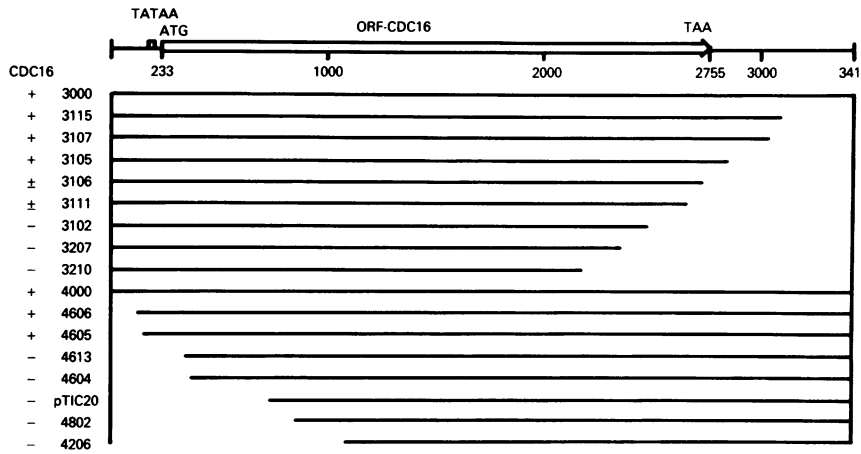


FIG. 2. Deletions inserted into M13 were recloned into pTI15 using the common SstI and SalI sites. Clones numbered 3000 and 4000 represent the original fragment in pTIC21 cloned into mp18 and mp19, respectively. C-terminal deletions 3115, 3107, and 3105 carry the entire CDC-ORF, as well as possible termination signals, and have full cdc16-1 complementing activity. Fragments 3106 and 3111 were deletions which are missing C-terminal polar residues in the coding sequence. The plasmids carrying these deletions were still able to weakly complement the cdc16-1 mutation. Other deletions missing more of the C-terminal region showed no CDC16 activity. The N-terminal deletion 4606 starts within the C-terminal coding sequence of the MAK11 gene and has cdc16-1 complementing activity. The deletion 4605 is only 69 bp from the start of the CDC16 ORF and still has full cdc16-1 complementing activity. Deletions 4613 and smaller ones lack the promoter region and a portion of the N-terminal coding sequence. Consequently, they had no cdc16-1 complementing activity.

residues, whose molecular weight and pI are 94967 and 6.60, respectively, assuming no protein modification. This ORF assignment for the CDC16 gene is consistent with the results of complementation tests of deletion plasmids summarized in Fig 2. The ORF of the MAK11 gene ends at bp 137 with (TA)A. Therefore, the space between these two ORFs is remarkably short, only 95 bp. Furthermore, one of the deletions, M4605, carries only 69 bp of sequence upstream from the start of the ORF and still is capable of complementing the cdc16-1 mutation. Between the two ORFs, there are possible termination signals for the MAK11 message, TAG...TATGT...TTT (11). Overlapping with these sequences are two TATA boxes in tandem, GTATAA/TATAAAA. As indicated in Fig. 3 (positions 177 to 212), this AT-rich region, possibly responsible for both MAK11 transcription termination and CDC16 initiation, has dyad symmetry over 32 bp except for the three G residues. Five bases preceding the start of the CDC16 ORF, there is another ATG codon. This frame is terminated, after only 2

amino acids, by TGA. This ATG codon probably reduces the level of CDC16 gene expression (12).

Within the coding sequence, there are four direct repeats of C(T/C)AATT (II) starting at position 594, resulting in an alternating $(Asn\ Ser)_4$ sequence. Another directly repeated sequence of 16 bp with one mismatch (III) is found starting at position 851. This produces a region in which 9 of 10 amino acid residues are Thr. Also there is a common sequence, CTCAATAAAAAGA (I), starting at base 78 and at base 2579 which resembles the possible CAAT sequence (13) in the MAK11 gene (T. Icho and R. B. Wickner, submitted). However, the significance of these sequences is unknown. The ORF ends at base pair 2755 with (TA)A. We can also identify possible termination signals, TAG...TAGT...TTT (11) and three AATAAAs, including the one at the termination codon. A fourth repeat sequence (IV), after the termination signals, presumably belongs to the neighboring gene.

The CDC16 protein contains a total of 16 cysteine residues, 5 of them clustered at the N-terminal, 4 placed about at the middle of the protein, and 3 located with a histidine residue close to the C-terminal. Each of these three clusters could form a metal binding DNA binding domain (5, 14) (Fig. 3).

The CDC16 Gene Is Essential

The EcoRV fragment (bases 1131-1645) within the CDC16 ORF was replaced by the URA3 gene, the fragment containing the disrupted (15) CDC16 fragment was transformed into a diploid strain, YTI81/YTI36 (ura3 CDC16 MAK11/ura3 cdc16 mak11-1), and Ura^+ transformants were selected (Fig. 4). Among diploid strains disrupted correctly, as judged by Southern blotting, three were temperature-sensitive and five were wild type. The former are interpreted as disruption within the wild type CDC16 gene and the latter as that within the temperature-sensitive allele. Their spore forming efficiency was significantly reduced, especially in the temperature-sensitive diploids. This may be related to the requirement for CDC16 in meiotic spore formation (16). In both cases, only two of four tetrad spores survived and all these germinated spore clones were Ura^- among a total of 20 tetrads examined. This result confirmed that this gene is essential for growth of yeast as expected from the temperature-sensitive nature of the cdc16-1 mutation.

CDC16 Transcription

RNA from age-fractionated cells was prepared as described (17) and generously supplied by Mitchell Smith. A Northern blot of this RNA was hybridized with a CDC16 probe and with a probe for the constitutive gene PYK1 (pyruvate kinase). Cells of all cell cycle stages showed a 2.8 kb transcript (Fig. 5), and

normalization by comparison with PYK1 message showed that there is no cell cycle regulation of CDC16 transcription (Fig. 5).

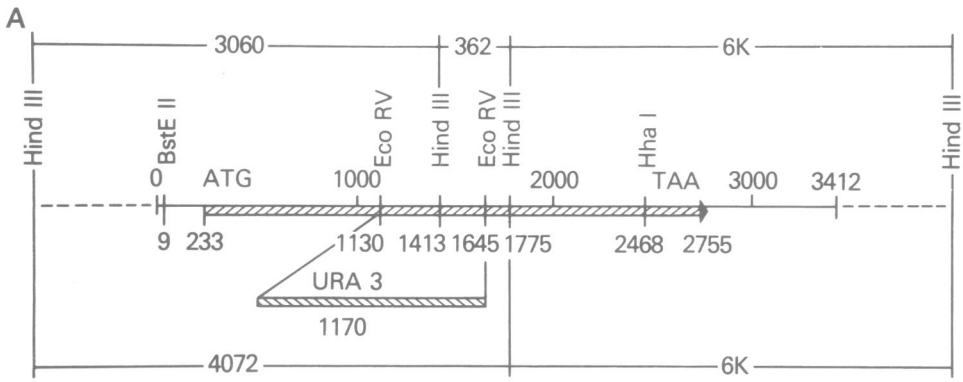
DISCUSSION

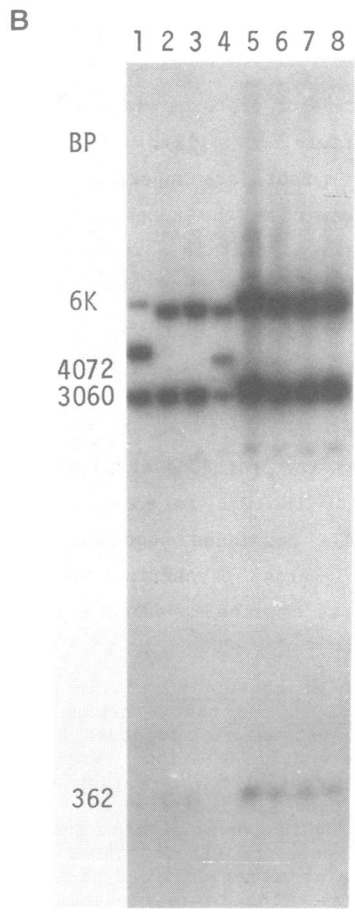
The CDC16 protein has a total of 16 cysteine residues. Twelve of them are clustered either at the N-terminal, in the middle, or close to the C-terminal of the sequence, forming metal-binding nucleic acid-binding domains originally proposed for TFIIIA (5), in which four cysteine or histidine residues are distributed in the pattern Cys X_{2-4} Cys X_{2-15} Cys/His X_{2-4} Cys/His (14). Since in the first cluster there are six cysteine residues and one histidine residue, two metals may possibly bind to this region. Alternatively, two different binding conformations may be possible. The first cluster is quite homologous to one of the TFIIIA repeats, while the second resembles a region in ferridoxin.

In yeast, several activator proteins which interact with upstream promoter regions have this structural motif, including GAL4, PPR1, ARGRII (18), and ADR1 (19). Since cdc16-1 mutants are reported to accumulate the predominant microtubule-associated protein at the nonpermissive temperature (20), it is possible that the CDC16 gene product is a repressor-like molecule binding DNA and regulates the expression of this microtubule-associated protein. Alternatively, this protein could be a structural component of the spindle that actually binds to the yeast chromosome. In any case, this protein must play an important role in the process of nuclear division.

The distance between the MAK11 and CDC16 genes is very short. The positive complementing activity of the deletion M4605 suggests that only 69 bp upstream from the start of the ORF is enough for the CDC16 activity. We identified two tandem TATA consensus sequences in this region. Northern analysis using T3 RNA polymerase transcripts of the EcoRV fragment in the middle of the CDC16 gene as a probe revealed a low abundance message of 2.8

FIG. 3. The DNA sequence of the CDC16 gene in pTIC21. The DNA sequence is shown together with the CDC16 and the C-terminal part of the MAK11 protein sequences. A small ORF, just before the start of the CDC16 coding sequence, is shown italicized. The end of deletions shown in Fig. 2 is marked by the arrows together with their cdc16-1 complementing activity in parenthesis. Two TATA boxes for the CDC16 gene are indicated by *s. The possible termination signals for the MAK11 and CDC16 genes, AATAAA, and TAG...TATGT...TTT (11), are italicized. One repeated sequence (I) and three tandem direct repeats, (II), (III), and (IV), as well as a region of dyad symmetry around the promoter region, are shown by lines. The three clusters of cysteine and histidine residues, which could possibly form metal-binding, nucleic acid-binding domains, are in the boxes.


A
3060
362
6K
Hind III
BstE II
Eco RV
Hind III
Eco RV
Hind III
Hha I
Hind III
0
ATG
1000
2000
TAA
3000
3412
9
233
1130
1413
1645
1775
2468
2755
URA 3
1170
4072
6K


B
1 2 3 4 5 6 7 8
BP
6K
4072
3060
362

kb, whose amount does not change during the cell cycle compared with PYK1 message. Consequently, we suggest that the CDC16 message is expressed constitutively.

It has been proposed that a constitutive promoter in yeast consists of an upstream poly(dA-dT) sequence (21) that allows the entry of RNA polymerase. In the case of the CDC16 promoter, there is no room left for such a sequence. If a separate entry site is really necessary for the initiation of transcription, we could assume that the DNA template may be already open, because of MAK11 transcription, and even that the same RNA polymerase molecule might be used to initiate CDC16 transcription. However, since the deletion M4605 is enough to complement the cdc16-1 mutation, this may not be the case. Instead, perhaps only the AT-rich sequence including the two TATAs is actually enough for the initiation of the moderate constitutive transcription observed. However, we have not yet completely eliminated the possibility that an artificial outside promoter activity originating from the vector may affect the expression of the cloned CDC16 gene.

There is an ATG codon just before the start of the CDC16 ORF that terminates just within the CDC16 ORF after only two amino acids. This would be expected to reduce the level of expression of CDC16 (12) if the mRNA starts before this upstream ATG codon. Since the mRNA start site is not yet known, it remains possible that translation of the CDC16 ORF begins at the second ATG at position 347. Other factors reducing the level of expression of CDC16 are the fact that less abundant tRNA codons (22) are frequently used in the CDC16 coding sequence and the dissimilarity of the sequence around the AUG to the

FIG. 4. Disruption of the CDC16 gene. (A) Construction. The SstI-SalI fragment containing the CDC16 gene in the plasmid pTIC21 was inserted into the corresponding sites in plasmid pTIC13. The 515 bp EcoRV fragment in this plasmid, pTIC30, was replaced by an 1170 bp HindIII fragment containing the URA3 gene, which was repaired by the Klenow fragment of DNA polymerase I. From this plasmid, pTI31, a BstEII-HhaI fragment containing the URA3-disrupted CDC16 gene was purified and transformed into a diploid strain, YTI81 (ura3 CDC16 MAK11 K_1^+)/YTI36 (ura3 cdc16-1 mak11-1). (B) Verification of the disruption. The HindIII digests of genomic DNAs from URA3$^+$ diploid transformants and their spore clones were tested for the integration of the URA3 gene by Southern hybridization (28, 29) using the nick-translated SstI-SalI fragment from pTIC21 as a probe. The replacement of two HindIII fragments of 3060 bp and 362 bp by a 4072 bp fragment indicates the correct disruption. The HindIII site on the left side in panel A is in plasmid pTIC19, which was previously sequenced. The exact location of the right HindIII site is not known. The weak hybridization band of about 2 kb seems to be derived from an homologous sequence in the yeast genome. (1) Plasmid pTIC1; (2) strain 1385; (3) YTI81 x YTI36; (4) a diploid Ura$^+$ transformant, D3. Two pairs of Ura$^+$ spore clones derived from D3: (5) 1A; (6) 1B; (7) 3A; and (8) 3B.

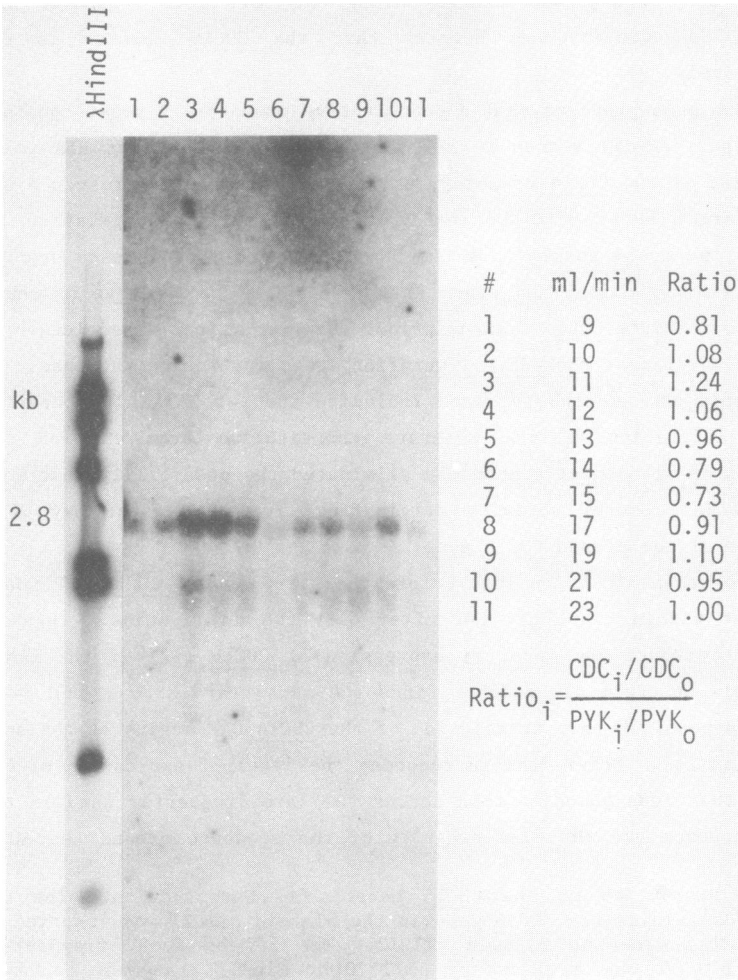


FIG. 5. The CDC16 transcript is not regulated by the cell cycle. The total RNAs from size-fractionated cells using an elutriator centrifuge were run on a glyoxal gel, transferred to a nitrocellulose filter, and hybridized with the T3 RNA polymerase transcript from plasmid pTIC56 (shown in figure). Plasmid pTIC56 carries the EcoRV fragment (base pairs 1130 to 1645) from pTIC21 inserted into the SmaI site of pT3/T7-18 (Bethesda Research Laboratories). These in vitro transcripts are complementary to the CDC16 message. The T7 RNA polymerase transcript from a plasmid containing a fragment of the PYK1 gene was used as a control (data not shown). The relative amount of each message was quantitated by scanning the x-ray film. The amount of each message (CDC_i and PYK_i) was first normalized to that of unfractionated cells (CDC_o and PYK_o), and then the normalized ratio of the two messages was calculated for each fraction.

consensus sequence for highly expressed yeast genes (23). In fact, CDC16-lacZ gene fusions at amino acid 185 (C47) and at amino acid 84 (C51) of CDC16, when expressed on the CEN plasmid pTI15, showed 3.2 units and 1.35 units of ß-galactosidase activity, respectively, compared to, for example, MATα2-lacZ fusions on a CEN plasmid (100 units) (24) and ENO1-lacZ fusions on a CEN plasmid (85 units) (25). Although other CDC16-lacZ fusions might give different activities, these results suggest that the level of expression of CDC16 is low.

ACKNOWLEDGMENT

We thank Dr. Mitchell Smith for providing mRNA preparations from age-fractionated cells and for his comments on the manuscript.

REFERENCES

1. Pringle, J. R., and Hartwell, L. H. (1981) In Strathern, J. N., Jones, E. W., and Broach, Jr. (eds), The Molecular Biology of the Yeast Saccharomyces, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 97-124.
2. Hartwell, L. H. (1974) Bacteriol. Rev. 38, 164-198.
3. Byers, B., and Goetsch, L. (1974) Cold Spring Harbor Symp. Quant. Biol. 38, 123-131.
4. Hartwell, L. H., and Smith, D. (1985) Genetics 110, 381-395.
5. Miller, J., McLachlan, A. D., and Klug, A. (1985) EMBO J. 4, 1609-1614.
6. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) J. Bacteriol. 153, 163-168.
7. Legerski, R. J., Hodnett, J. L., and Gray, H. B., Jr. (1978) Nucleic Acids Res. 5, 1445-1464.
8. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33, 103-119.
9. Biggin, M. D., Gibson, T. J., and Hong, G. F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963-3965.
10. Messing, J. (1983) Methods Enzymol. 101, 20-78.
11. Zaret, K. S., and Sherman, F. (1982) Cell 28, 563-573.
12. Sherman, F., and Stewart, J. W. (1981) In Strathern, J. N., Jones, E. W., and Broach, Jr. (eds), The Molecular Biology of the Yeast Saccharomyces, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 301-333.
13. Dobson, M. J., Tuite, M. F., Roberts, N. A., Kingsman, A. J., and Kingsman, S. M. (1982) Nucleic Acids Res. 10, 2625-2637.
14. Berg, J. M. (1986) Science 232, 485-487.
15. Rothstein, R. J. (1983) Methods Enzymol. 101, 202-211.
16. Simchen, G. (1974) Genetics 76, 745-753.
17. Gordon, C. N., and Elliot, S. G. (1977) J. Bacteriol. 129, 97-100.
18. Messenguy, F., Dubois, E., and Descamps, F. (1986) Eur. J. Biochem. 157, 77-81.
19. Hartshorne, T. A., Blumberg, H., and Young, E. T. (1986) Nature (London) 320, 283-287.
20. Shriver, K., and Byers, B. (1977) J. Cell Biol. 75, 297a.
21. Struhl, K. (1985) Proc. Natl. Acad. Sci. USA 82, 8419-8423.
22. Ikemura, T. (1982) J. Mol. Biol. 158, 573-597.
23. Hamilton, R., Watanabe, C. K., and de Boer, H. A. (1987) Nucleic Acids Res. 15, 3581-3593.

24. Hall, M. N., Hereford, L., and Herskowitz, I. (1984) Cell 36, 1057-1065.
25. Uemura, H., Shiba, T., Machida, M., Natsui, I., Jigami, Y., and Tanaka, H. (1987) J. Biochem. 102, 181-189.
26. Icho, T., Bulawa, C. E., and Raetz, C. R. H. (1985) J. Biol. Chem. 260, 12092-12098.
27. Meinkoth, J., and Wahl, G. (1984) Anal. Biochem. 138, 267-284.
28. Smith, G. E., and Summers, M.D. (1980) Anal. Biochem. 109, 123-129.