MOLECULAR AND CELLULAR BIOLOGY, June 1991, p. 3317–3325
0270-7306/91/063317-09$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 11, No. 6

The Human *CCG1* Gene, Essential for Progression of the G_1 Phase, Encodes a 210-Kilodalton Nuclear DNA-Binding Protein

TAKESHI SEKIGUCHI,[1] YUKIKO NOHIRO,[1] YASUHARA NAKAMURA,[2] NAOKI HISAMOTO,[1] AND TAKEHARU NISHIMOTO[1]*

Department of Molecular Biology, Graduate School of Medical Science, Kyushu University, Maidashi, Fukuoka 812,[1] and Laboratory of Biological Chemistry, School of Pharmaceutical Science, Showa University, Shinagawa-ku, Tokyo 142,[2] Japan

Received 22 October 1990/Accepted 23 March 1991

The human *CCG1* gene complements *ts*BN462, a temperature-sensitive G_1 mutant of the BHK21 cell line. The previously cloned cDNA turned out to be a truncated form of the actual *CCG1* cDNA. The newly cloned *CCG1* cDNA was 6.0 kb and encoded a protein with a molecular mass of 210 kDa. Using an antibody to a predicted peptide from the CCG1 protein, a protein with a molecular mass of over 200 kDa was identified in human, monkey, and hamster cell lines. In the newly defined C-terminal region, an acidic domain was found. It contained four consensus target sequences for casein kinase II and was phosphorylated by this enzyme in vitro. However, this C-terminal region was not required to complement *ts*BN462 mutation since the region encoding the C-terminal part was frequently missing in complemented clones derived by DNA-mediated gene transfer. CCG1 contains a sequence similar to the putative DNA-binding domain of HMG1 in addition to the previously detected amino acid sequences common in nuclear proteins, such as a proline cluster and a nuclear translocation signal. Consistent with these predictions, CCG1 was present in nuclei, possessed DNA-binding activity, and was eluted with similar concentrations of salt, 0.3 to 0.4 M NaCl either from isolated nuclei or from a DNA-cellulose column.

The cell cycle of eucaryotic cells is composed of the consecutive phases G_1, S, G_2, and M. In the G_1 phase, depending on the stimulus of growth factors, protein and RNA molecules are produced and the new cell cycle initiates (25). Unresolved problems concerning the start of the cell cycle include how the stimulus of external growth factors reaches the nucleus and how cells recognize the accumulation of materials required to enter the new cell cycle. To investigate the initiation of the cell cycle at the molecular level, it is essential to identify the genes involved. In both budding and fission yeasts, various temperature-sensitive (*ts*) mutants defective in the progression of the cell cycle have been isolated and have proven to be most useful for identifying the genes required for cell cycle progression (12, 22). Various ts^+ cell cycle mutants have also been isolated from cultured animal cells (3). The FT210 cell line, a *ts* G_2 mutant of FM3A, has a *ts cdc2* gene product; thus, the *cdc2* gene is apparently essential for the cell cycle of animal cells (36).

We isolated *ts* cell cycle mutants from the BHK21 cell line derived from the Syrian hamster and then classified them into complementation groups (21). One of these *ts* mutants, *ts*BN462, has a defect in progression of the G_1 phase, which is that after release from the G_1 block, it is unable to enter the S phase at a nonpermissive temperature. But once it enters the S phase, DNA replication progresses normally. By using this mutant as the recipient of the DNA-mediated gene transfer, a human gene located between q11 and q13 on the X chromosome has been cloned and designated the cell cycle gene 1 (*CCG1*) (5, 31, 32). Its cDNA complements *ts*13 in addition to *ts*BN462, both of which were independently isolated from BHK21 cells and belong to the same complementation group (31). The *ts*13 cell line has been investigated extensively by Liu et al. (16). After release from the G_1 block, it cannot enter the S phase at a nonpermissive temperature and the expression of both the thymidine kinase and the dihydrofolate reductase genes is markedly reduced. The execution point of *ts*13 mutation is 3 h before the S phase.

From ts^+ transformants of the *ts*BN462 cell line, two groups of human DNAs were cloned. One was carried by all ts^+ transformants of *ts*BN462 and encoded the *CCG1* cDNA complementing the *ts*BN462 mutation. The other group was carried in half the number of ts^+ transformants, through DNA-mediated gene transfer. We had considered that it was not related to the *CCG1* gene; however, in this research, both genomic DNAs were found to derive from human *CCG1* genomic DNA, and the previously cloned *CCG1* cDNA turned out to be a truncated form of the actual one. The cloned *CCG1* cDNA encoded a protein with a molecular mass of 210 kDa instead of the previously estimated mass of 180 kDa. CCG1 has DNA-binding activity and is present in the nuclei. Interestingly, the C-terminal part of CCG1, which we found during this research, contains the target consensus sequence for casein kinase II and was phosphorylated by this enzyme in vitro.

MATERIALS AND METHODS

Cell line and cell culture. The *ts*BN462 cell line is a *ts* mutant of the BHK21/13 cell line (21). The KB cell line is derived from a human laryngeal carcinoma. The K-1, K-6, and K-7 cell lines are primary ts^+ transformants of *ts*BN462 transfected with human DNA, and the K-1-1 and K-1-2 cell lines are secondary ts^+ transformants of the K-1 cell line (32). The K-1-2-1, K-1-2-2, and K-1-2-3 cell lines are tertiary ts^+ transformants of the K-1-2 cell line. The COS-7 cell line

* Corresponding author.

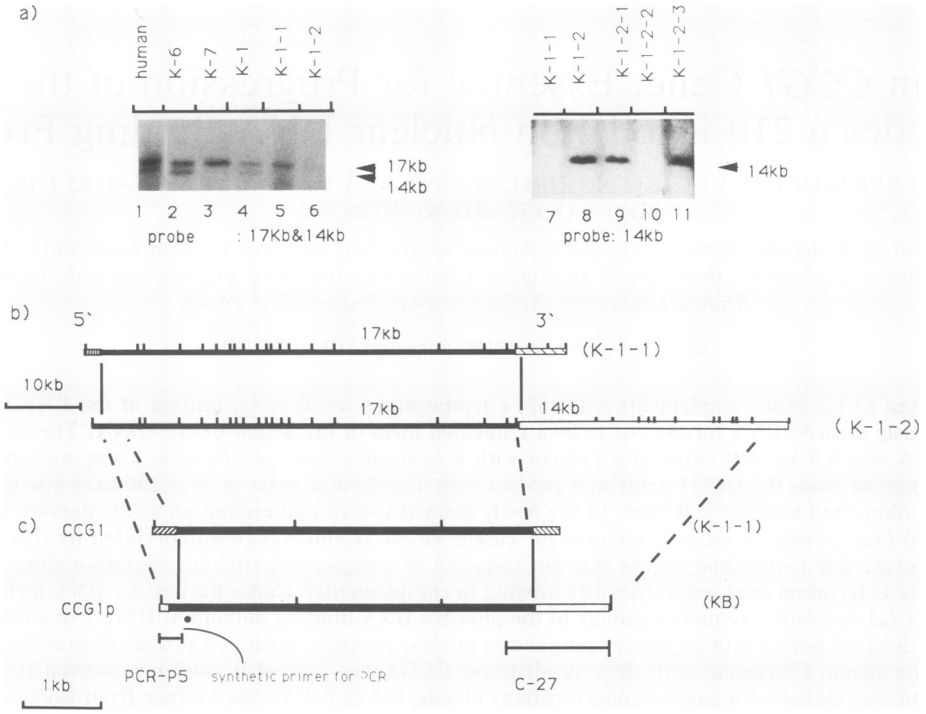

FIG. 1. Human *CCG1* gene isolated from the secondary ts^+ transformants of *ts*BN462, K-1-1 and K-1-2. (a) Conservation of original *Eco*RI human DNA fragments (17 and 14 kb) through ts^+ transformation. DNAs (20 μg) prepared from original KB cells, the primary ts^+ transformants (K-1, K-6, and K-7), the secondary ts^+ transformants (K-1-1 and K-1-2), and the tertiary ts^+ transformants (K-1-2-1, K-1-2-2, and K-1-2-3) were digested with *Eco*RI, electrophoresed in a 1% agarose gel, and then analyzed, using as a probe the Alu-free regions of 17-kb *Eco*RI (lanes 1 through 6) and 14-kb *Eco*RI fragments (lanes 1 through 11) as described previously (32). Arrowheads indicate the positions of the *Eco*RI 17- and 14-kb fragments. (b) *Eco*RI restriction map of the genomic DNAs. The upper and lower lines show the genomic DNA isolated from the K-1-1 and the K-1-2 cell lines, respectively. The vertical black lines indicate the region shared by both genomic DNAs. (c) *Eco*RI restriction map of the *CCG1* and *CCG1*p cDNA. The upper and lower lines indicate the *CCG1* and the *CCG1*p cDNA, respectively. *CCG1*p represents the newly cloned cDNA (see text). The dense black lines indicate the region common to both cDNAs. The dotted lines indicate the region newly found in the *CCG1*p cDNA.

is the simian virus 40 T-antigen-positive transformant of the Green Monkey CV-1 cell line (11).

All cell lines were cultured in Dulbecco's modified Eagle medium containing 10% calf serum, penicillin (100 U/ml), and streptomycin (100 μg/ml) in a humidified atmosphere of 10% CO_2–90% air. The *ts*BN462 cell line was maintained at 33.5°C, while the other cell lines were kept at 37.5°C. Transformation of the *ts*BN462 cell lines was performed by the calcium phosphate precipitation method, as described previously (32).

Isolation of nucleic acid and filter hybridization. High-molecular-weight DNA was isolated from cultured cells by phenol extraction of the sodium dodecyl sulfate (SDS)-disrupted cells. Plasmid and λ phage DNA were isolated by the procedure of Maniatis et al. (19). The poly(A)$^+$ RNA was purified by oligo(dT) chromatography. Southern and Northern (RNA) blot hybridizations were performed as described previously (23, 31). Restriction endonucleases were purchased from Takara Shuzo (Kyoto, Japan), and digestion was carried out under the conditions recommended by the supplier.

Construction and screening of a genomic DNA library. A genomic DNA library was constructed essentially as described previously (32). As vector, λ phage Charon 4A and cosmid pTAK1 were used. All cloning experiments were carried out according to the Guidelines for Recombinant DNA Research issued by the Ministry of Education, Science and Culture of Japan.

Construction of the cDNA library. The cDNA library of the poly(A)$^+$ RNA of KB cells was constructed by using the Amersham cDNA and λgt10 cDNA cloning system (Amersham Corp., Arlington Heights, Ill.) and was screened as described previously (31).

5′-end amplification of cDNAs. One microgram of poly(A)$^+$ RNA was reversely transcribed to the cDNA by using a random primer. The cDNA was then tailed by terminal deoxynucleotidyl transferase as described previously (10). The poly(A)-tailed cDNA was amplified by polymerase chain reaction (26) as follows. The mixture of cDNA, deoxyribosylthymine 16 adapter (10 pmol) GATATCTGGC CAGCGGCCGCTTTTTTTTTTTTTTTT, adapter (25 pmol) GATATCTGGCCAGCGGCCGCT, and amplification primers (ccg718, 25 pmol) GGAACTACTGAAGTCCACTTTTC was denatured (5 min at 95°C) and cooled to 72°C. Following the addition of 2.5 U of *Taq* DNA polymerase (Perkin-Elmer-Cetus), the mixture was overlaid with 30 μl of mineral oil (Sigma) at 72°C and annealed at 50°C for 2 min. The cDNA was then extended at 72°C for 3 min. Forty cycles of amplification were carried out by a step program (94°C, 1 min; 50°C, 2 min; 72°C, 3 min), and, finally, 15-min extension was performed at 72°C.

Sequencing and computer analysis. The DNA clones were

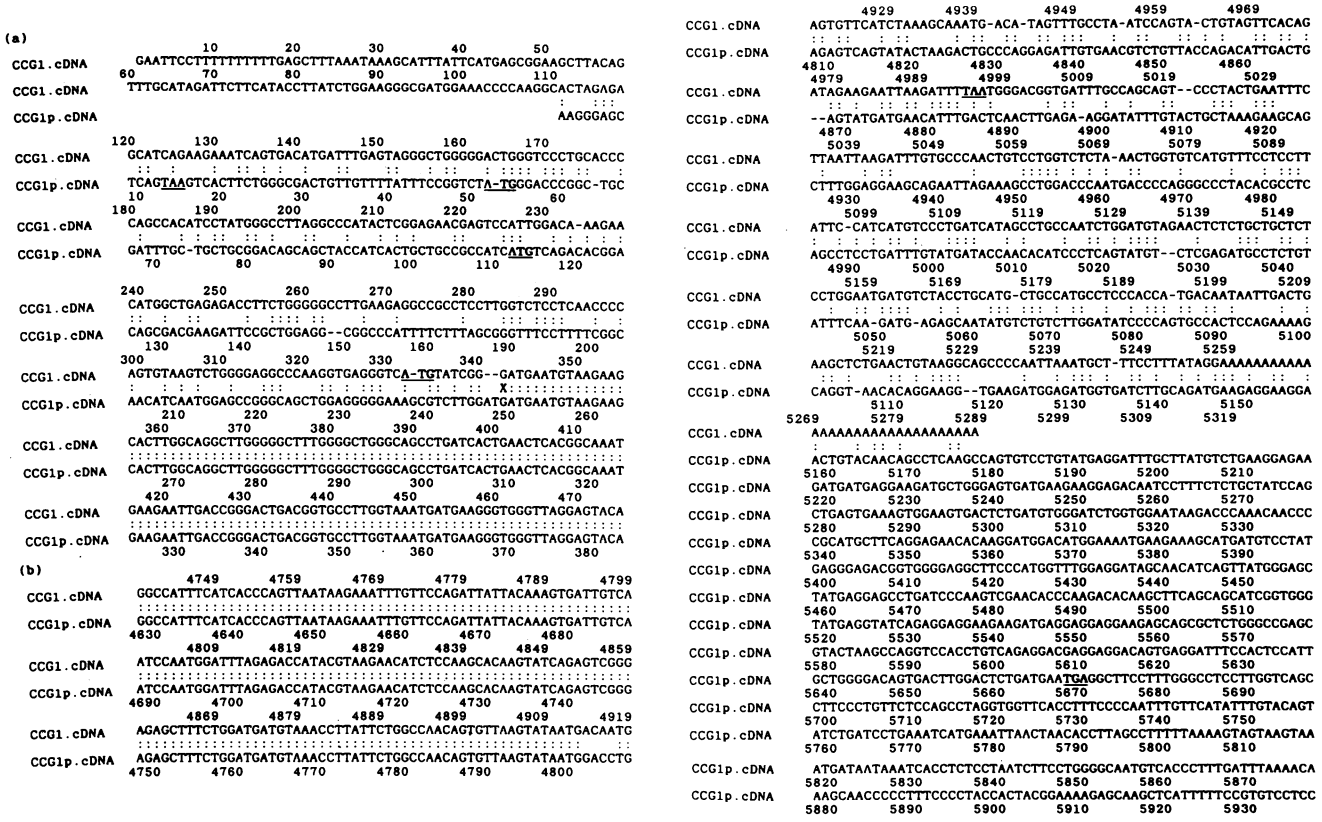

FIG. 2. Comparison of the nucleotide sequence between the *CCG1* and *CCG1*p cDNA. (a) Comparison at the 5′ terminus. The nucleotide sequence in the region downstream from the 249th nucleotide of *CCG1*p is shown. X is the same for both cDNAs. Underlining indicates the position of the stop codon (TAA) and the start codon (ATG). (b) Comparison at the 3′ terminus. The *CCG1* cDNA ends at the 5177th base of the *CCG1*p. Underlining indicates the position of the stop codon. The homologous bases are shown by colons.

prepared for nucleotide sequencing with a Takara deletion kit as recommended by the supplier, and they were then sequenced in both directions by the dideoxy chain termination method (28). The determined nucleotide sequences were compiled and analyzed with the programs in DNASIS (Hitachi Software Engineering Co., Ltd., Yokohama, Japan).

Preparation of antibody. The synthetic peptides were prepared by Hoechst Japan Co., according to the sequence of *CCG1* cDNA (Saitama, Japan). Cysteine was added to the carboxy-terminal ends of the peptides in order to couple the carrier protein keyhole limpet hemocyanin. Synthetic peptides coupled to the carrier protein were used to immunize a rabbit, as described previously (30).

Expression and detection of a part of the CCG1 protein in *Escherichia coli*. The 1733-bp *Bcl*I-*Hin*dIII fragment of the previously isolated *CCG1* cDNA (31) was inserted into the *Bam*HI-*Hin*dIII-digested pUR292 (27) and then transformed into *E. coli* F′11 *recA*. *E. coli* F′11 *recA* containing the fusion plasmid and pUR292 was grown in Luria-Bertani broth at 37°C to A_{660} = 0.6, and then isopropyl-β-D-thiogalactopyranoside (IPTG) was added (final concentration 0.5 mM). The incubation was allowed to proceed for another 3 h. The products were analyzed by SDS–10% polyacrylamide gel electrophoresis (PAGE) and were then visualized by Coomassie blue staining.

Immunoblotting. Cells were lysed in buffer containing 62.5 mM Tris-HCl (pH 6.8), 10 mM 2-mercaptoethanol, 3% (wt/vol) SDS, and 20% glycerol. Cellular proteins were electrophoresed in an 8 or 10% SDS-polyacrylamide slab gel and analyzed by immunoblotting, as described previously (6), by using the antibody against the synthetic polypeptide and a vector stain ABC kit. Protein concentration was determined by Bradford's method (4).

Preparation of nuclei. Cells were collected and fractionated into nuclei and cytoplasm as described elsewhere (24). The isolated nuclei were suspended in ice-cold hypotonic buffer and then kept on ice. All procedures for cell fractionation were performed with freshly prepared nuclei at a concentration of 10^8 nuclei per ml. A hypotonic buffer containing the desired salt concentration was used for extraction with the salt. Nuclear supernatant and residual nuclei were separated by centrifugation. Each fraction was analyzed by immunoblotting with the anti-CCG1 antibody.

DNA-binding assay. The isolated nuclei were suspended in buffer A, which contained 10 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, 5 μg each of aprotinin, pepstatin, and leupeptin per ml, and 1 mM phenylmethylsulfonyl fluoride. The isolated nuclei were suspended in NaCl at a final concentration of 0.4 M. The salt-extracted proteins and the residual proteins were separated by centrifugation. After dilution with buffer B, which contain buffer A and 1 mM dithiothreitol, the salt-extracted protein was loaded onto a double-stranded calf thymus DNA-cellulose column. Bound proteins were then eluted with buffer containing 10 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, and increasing concentrations of NaCl. The flowthrough and the eluted fractions were precipitated

FIG. 3. Predicted amino acid sequence of CCG1. On the basis of the nucleotide sequence of *CCG1*p, the amino acid sequence was deduced. Positions of the synthetic peptide, the HMG1 box, the nuclear location signal, and the repeated domain are indicated. Boldface shows the amino acid sequence similar to the DNA-binding motif of histone [S(T)PXX], the nuclear location signal (PPKKKRRV), and the consensus target sequence of the protein kinase casein kinase II [S-(E/D)n], cAMP-dependent kinase (RRXSX), and Ca^{2+}/calmodulin-dependent kinase (RXYS/T).

by 50% ammonium sulfate and analyzed by electrophoresis on 8% SDS-polyacrylamide gels and by immunoblotting.

Indirect immunofluorescence staining. Cells grown on glass coverslips were fixed in 3% (wt/vol) paraformaldehyde in phosphate-buffered saline (PBS), permeabilized with 1% Nonidet P-40 in PBS containing glycine, and stained sequentially with the antipeptide and then with the rhodamine-conjugated goat anti-rabbit immunoglobulin G (Tago Inc., Burlingame, Calif.), as described previously (24). DNA was stained with Hoechst dye 33258 (1 μg/ml) (Hoechst Japan, Ltd.).

Phosphorylation by casein kinase II. The 1.0-kb *Acc*I-*Dra*I cDNA fragment, which encodes the region of amino acids 1591 to 1872 in the CCG1 containing the target sequence to the casein kinase II, was subcloned into the *E. coli* expression vector pUR290. The resulting clone was transfected to *E. coli* XL1-blue strain. The LacZ fusion protein was overproduced by the addition of IPTG and was then purified by Protosorb LacZ immunoaffinity adsorbent column (Promega Corp., Madison, Wis.) as recommended by the supplier. This was then subjected to phosphorylation by casein kinase II, as described elsewhere (33).

RESULTS

Isolation of human genomic DNA present in K-1-2 but not in K-1-1 cells, the secondary transformant of *ts*BN462 cells. Among the ts^+ transformants of *ts*BN462, the 14-kb *Eco*RI fragment was found in one series of ts^+ transformants, including K-1-2 cells, but not in the other transformants, K-7 (primary transformant), K-1-1 (secondary transformant), and K-1-2-2 (tertiary transformant) (Fig. 1a), indicating that this fragment was not required to complement *ts*BN462 mutation. We therefore isolated human *CCG1* genomic DNA containing only the conserved human DNA from the K-1-1. Two of the tertiary ts^+ transformants derived from the secondary ts^+ transformant, K-1-2 cells, however, still contained the 14-kb *Eco*RI fragment (Fig. 1a). We therefore thought that this DNA fragment was probably tightly linked to the human *CCG1* gene, though it was not required to complement the *ts*BN462 mutation. To clarify the correlation between the cloned human *CCG1* gene and the 14-kb *Eco*RI fragment, we isolated the human DNA from the K-1-2 cells by using the human-specific Alu sequence as a probe.

The obtained human DNA was about 120 kb, and the 5′ region of 70 kb had the same restriction map as the previously cloned *CCG1* genomic DNA (Fig. 1b), indicating that both have the same base sequence. The rest of the newly cloned human DNA, that is, the 3′ region of 50 kb, contained the 14-kb *Eco*RI fragment unique to the K-1-2 cell line as shown in Fig. 1b, K-1-2. Thus, the previously cloned *CCG1* genomic DNA was disrupted at the 5′ end of 14-kb *Eco*RI fragment from the human DNA isolated in this research (Fig. 1b).

Isolation of the complete *CCG1* cDNA (*CCG1*p) from human cells. To investigate whether the 14-kb *Eco*RI fragment of human DNA is part of the actual *CCG1* gene, a *CCG1* cDNA was newly isolated from human cells (Fig. 1c). We designated the newly cloned *CCG1* cDNA as the CCG1p (p = perfect), to distinguish it from the previous one.

The cDNA clone C-27 was obtained from a human cDNA library and was found by Southern blot analysis to be transcribed not only from the 3′ part of previously isolated *CCG1* genomic DNA but also from the DNA fragment unique to the K-1-2 secondary ts^+ transformant (Fig. 1c). The nucleotide sequence of C-27 is common to those of *CCG1* and *CCG1*p cDNAs up to the 4801st base of *CCG1*p, but there is no similarity from the 4802nd base downstream. The open reading frame of *CCG1*p continues until the 5667th base (Fig. 2b). Thus, the 3′-terminal region of previously cloned *CCG1* cDNA was not transcribed from the human *CCG1* gene.

Since the previously cloned *CCG1* cDNA was isolated from the cDNA library of the secondary ts^+ transformant of *ts*BN462, it should contain cDNAs transcribed from the hamster DNA flanking the integrated human gene. Therefore, the 5′ part of previous *CCG1* cDNA may also be transcribed from hamster DNA. To resolve this question, we carried out 5′-end amplification of the human cDNA library by using an internal sequence of the *CCG1* cDNA as a primer as described in Materials and Methods. The nucleotide sequence of the cloned cDNA is shown as *CCG1*p in Fig. 2a. At the 5′-proximal end, there is no similarity between *CCG1*p and *CCG1* cDNA, though both cDNAs have the same sequence in an internal region. Consistent with this, the 5′-proximal 340 bp of *CCG1* cDNA nonhomologous to those of *CCG1*p did not hybridize with human genomic DNA, but the corresponding 5′ end of *CCG1*p cDNA did hybridize with human DNA (data not shown).

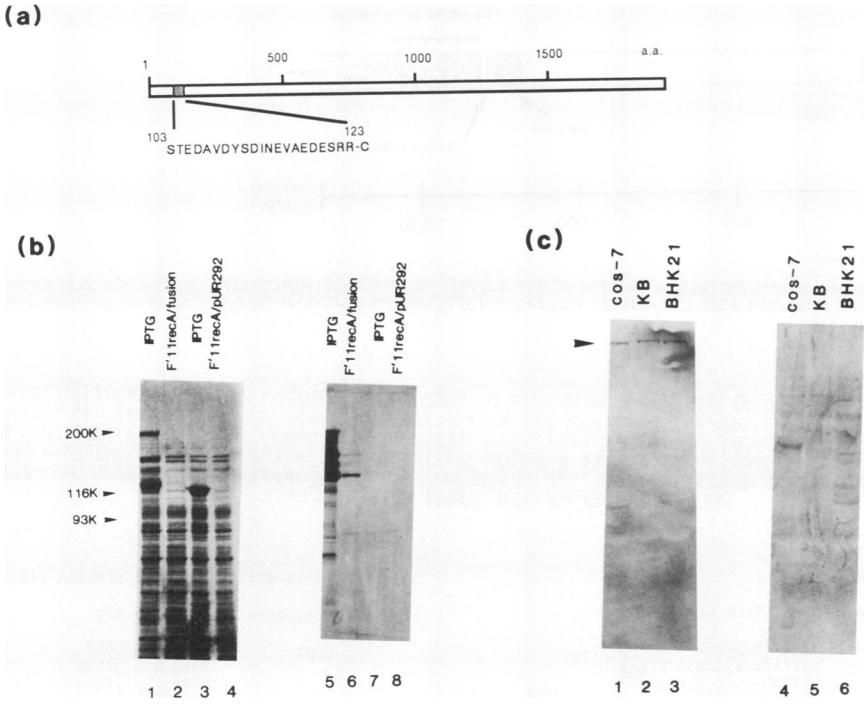

FIG. 4. Immunoblot analysis of CCG1. (a) Location of the synthetic peptide in the CCG1 protein, which was used as the antigen to prepare the antibody, and its amino acid sequence. This region is common to both *CCG1* and *CCG1*p cDNA. (b) Identification of the CCG1-LacZ fusion protein expressed in *E. coli* and its antigenicity to anti-CCG1. The CCG1-LacZ fusion protein was constructed as described in Materials and Methods. *E. coli*-extracted proteins of 50 μg were analyzed by SDS-PAGE (lanes 1 through 4), transferred to a nitrocellulose filter, and then immunoblotted with serum immunized against the synthetic peptide (lanes 5 through 8). Lanes: 1 and 5, the extract from *E. coli* F'11 *recA* cells harboring the fusion gene treated with IPTG; 2 and 6, the same cell line as in lane 1 without IPTG induction; 3 and 7, the extract from *E. coli* F'11 *recA* harboring the LacZ gene alone treated with IPTG; 4 and 8, the same cell line as in lane 3 without IPTG induction. (c) Presence of the CCG1 protein in various animal cell lines. Proteins of 50 μg extracted from cells were electrophoresed by SDS-PAGE, transferred to nitrocellulose filters, and then probed with immune serum alone (lanes 1 through 3) or with immune serum in the presence of the synthetic peptide (lanes 4 through 6). Lanes: 1 and 4, monkey COS-7 cells; 2 and 5, human KB cells; 3 and 6, hamster BHK21 cells.

Thus, the 5′ end of the previously cloned *CCG1* cDNA was not transcribed from human *CCG1* DNA.

Identification of CCG1 protein in vivo. In the CCG1 reading frame, there are two start codons, at the 52nd base and at the 112th base, and upstream of these start codons there is the stop codon TAA at the 13th base, indicating that both the ATGs could be the initiator codon of the CCG1 protein. Thus, CCG1 consists of either 1872 or 1852 amino acids, and so its molecular mass should be 212,666 or 210,724 Da, depending on the initiation site (Fig. 3). To identify the presence of CCG1 protein in the cells, an antibody to putative CCG1 protein was prepared by using an oligopeptide synthesized on the basis of the predicted amino acid sequence. The amino acid sequence of the peptide and its location in CCG1 are shown in Fig. 4a. The anti-peptide antibody specifically recognized a LacZ fusion protein containing part of the predicted CCG1 protein (Fig. 4b). In addition to the intact LacZ fusion protein of 200 kDa, many degradation products were recognized by the anti-CCG1 antibody as well.

A protein with a molecular mass of more than 200 kDa, consistent with the predicted molecular mass of CCG1, was recognized in extracts of human KB, monkey COS-7, and hamster BHK21 cell lines by immunoblot analysis (Fig. 4c). These bands disappeared in the presence of a synthetic peptide as a competitor. Thus, the bands of over 200 kDa correspond to CCG1, which is well conserved between species.

Structure of CCG1 protein. On the basis of the predicted amino acid sequence, several characteristic domains of CCG1 were found. In the N-terminal region, there are two proline clusters which have been suggested to be responsible either for the nuclear localization or for the stability of proteins belonging to the transcription factor family (20, 29, 37).

In the C-terminal region, we previously noted the presence of a nuclear location signal and a repeated structure of about 120 amino acids (31). We have now identified a sequence similar to the putative DNA-binding domain of HMG1 and a domain consisting of acidic amino acids (Fig. 5).

In the previously cloned cDNA, the repeated structure was disrupted at the 3′-proximal end of the 2nd repeat, that is, the 1584th amino acid of CCG1p. This domain turned out to consist of two full repeats and a third repeat consisting of only a few extra amino acids, as shown in Fig. 5b.

The HMG1-like domain is located just upstream of the nuclear location signal. A similar amino acid sequence has been found to be present three times in human upstream binding factor (hUBF), in which these sequences are responsible for DNA-binding activity of hUBF (14). When the amino acid sequence was compared with the HMG1 boxes of hUBF and with HMG1 itself, the HMG1-like domain of CCG1 was found to have significant similarity to the putative DNA-binding domain of HMG1. The same region of CCG1 has similarity with the promoter-binding domain of σ^{30} of

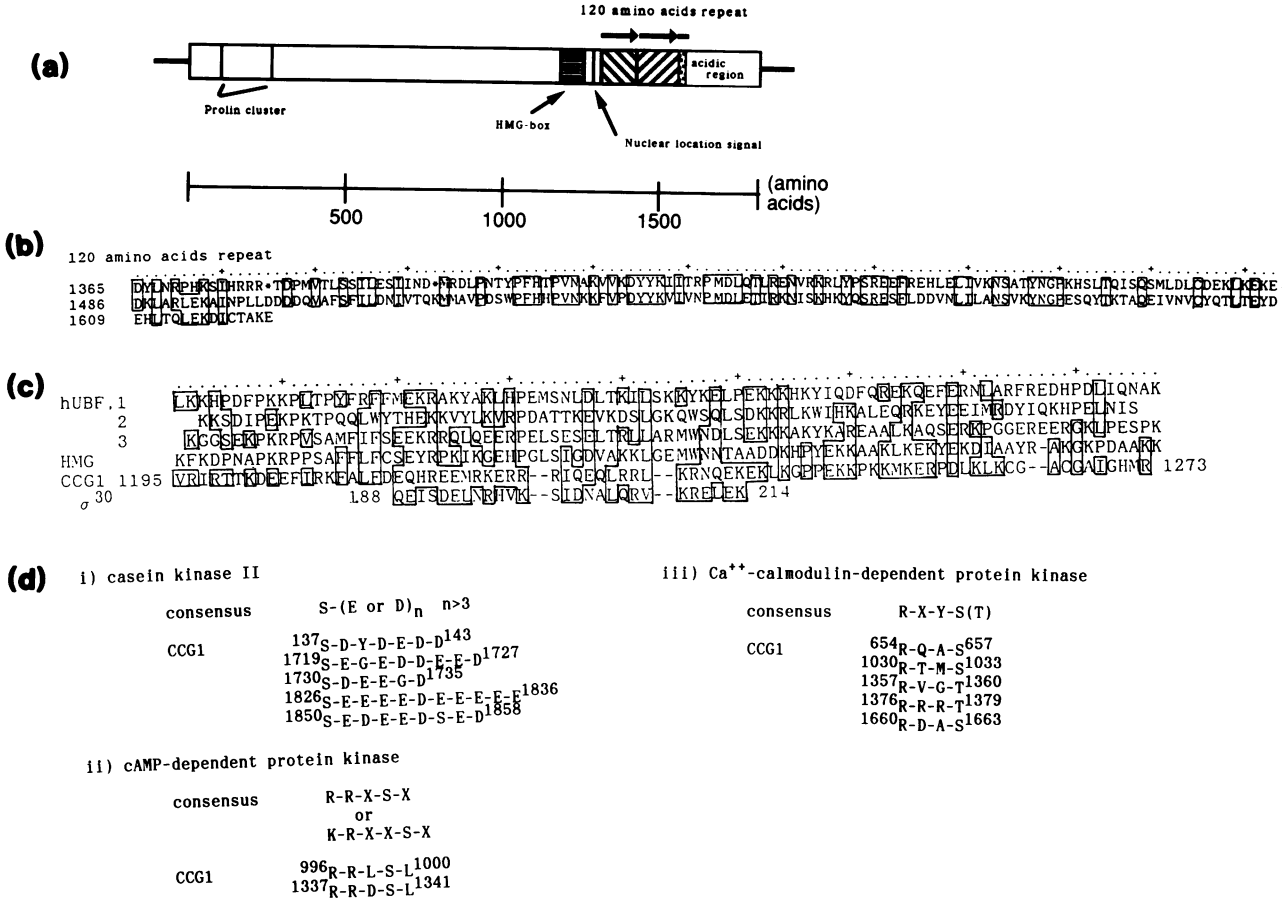

FIG. 5. Characteristic structure of CCG1 protein. (a) Schematic structure of CCG1 protein. The lines and regions in the boxes are explained above or below the related area. The lower line is a graduation in which the number of amino acids functions as the unit. (b) The amino acid sequence of the C-terminal repeated domain. The boxed regions indicate identical and chemically conserved amino acids. (c) The amino acid sequence of the domain similar to the putative DNA-binding domain of HMG1 and to the HMG1 box of hUBF. The amino acid sequence of three HMG1 boxes of hUBF, the putative DNA-binding domain of HMG1, and the region of σ^{30} were set in array and compared with the predicted HMG1 box of CCG1. The conserved amino acids were grouped as follows: A, S, T; I, L, M, V; E, D, N, Q; H, K, R; and F, W, Y. (d) Amino acid sequences found in CCG1 protein, similar to the consensus sequence phosphorylated with various protein kinases.

Bacillus subtilis (13) (Fig. 5c). This suggests that the CCG1 protein has DNA-binding activity. The presence of seven sequences similar to the DNA-binding motif of histone, S(T)PXX (34), also supports this notion (Fig. 3).

Like HMG1 and hUBF, CCG1 has an acidic tail which is encoded by the 3'-terminal region of the CCG1 cDNA cloned in this study. Of the terminal 248 amino acids, 35% of the amino acids were acidic. There are four S-(E or D)n sequences which are consensus target sequences for casein kinase II (Fig. 3 and 5d) (8). The region containing these sequences was specifically phosphorylated by casein kinase II in vitro (Fig. 6). Thus, native CCG1 may be modified by casein kinase II. The consensus target sequences for other protein kinases, such as cAMP-dependent and Ca^{2+}/calmodulin-dependent kinases, were also found in the CCG1 protein, suggesting that CCG1 may be multiply phosphorylated in vivo.

Nuclear localization and DNA-binding activity of the CCG1. On the basis of its nucleotide sequence, CCG1 appears to have amino acid sequences characteristic of nuclear protein, such as a proline cluster, an HMG1 homologous region, and a nuclear location signal similar to that of simian virus 40 (15, 32). By using the antipeptide antibody, we confirmed the nuclear location of the CCG1 protein.

BHK21 cells were doubly stained with the antibody and the DNA-staining dye, Hoechst 33258. As shown in Fig. 7, the same area that was stained with Hoechst 33258 was also stained with the antibody (Fig. 7a and c). In the presence of synthetic peptide, which was used to prepare the antibody, the nuclei were not stained (Fig. 7b and d), indicating that the nuclei were specifically stained with anti-CCG1 antibodies and showing that CCG1 is predominantly nuclear.

This was further confirmed by using isolated nuclei. BHK21 cells were fractionated into cytoplasmic and the nuclear fractions. CCG1 was released from nuclei by increasing ionic strength (Fig. 8). At 0.4 M NaCl, the CCG1 was extracted from the nuclei.

The DNA-binding activity of CCG1 was then investigated. A salt extract of the nuclei was diluted and then applied to a DNA-cellulose column, and the bound proteins were eluted by buffer containing increasing concentrations of NaCl. The eluted protein was concentrated by ammonium sulfate precipitation and analyzed by immunoblotting with anti-CCG1 antibody. As shown in Fig. 9, the CCG1 protein bound the

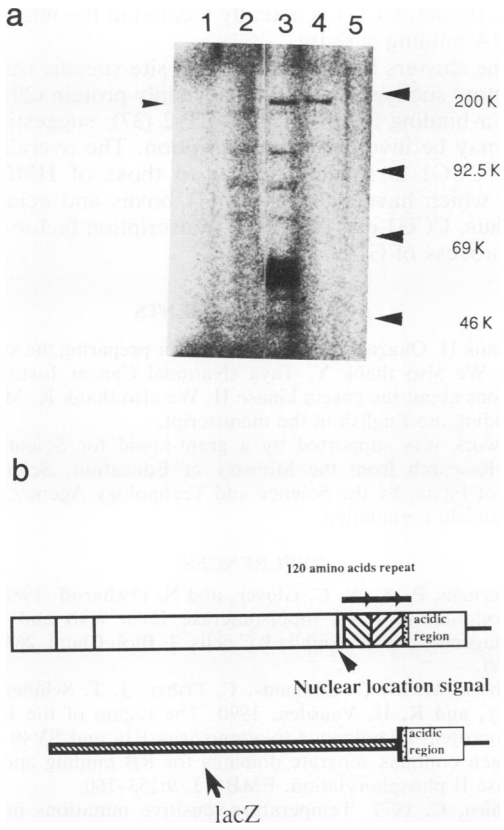

FIG. 6. Phosphorylation of the acidic region of CCG1 protein by casein kinase II in vitro. The 3′ acidic region of CCG1 protein was fused with LacZ protein (b). The fused protein was phosphorylated by casein kinase II as described in Materials and Methods. (a) Lanes: 1, extract of *E. coli* XL1-blue; 2, extract of *E. coli* XL1-blue containing pUR290; 3, extract of *E. coli* X11-blue expressing the fused protein; 4 and 5, affinity-purified LacZ-CCG1 fusion protein produced in *E. coli* XL-blue. Lanes 1 through 4 were treated with casein kinase II. Lane 5 was not treated with casein kinase II. The arrowhead to the left shows the position of phosphorylated LacZ-CCG1 fusion protein.

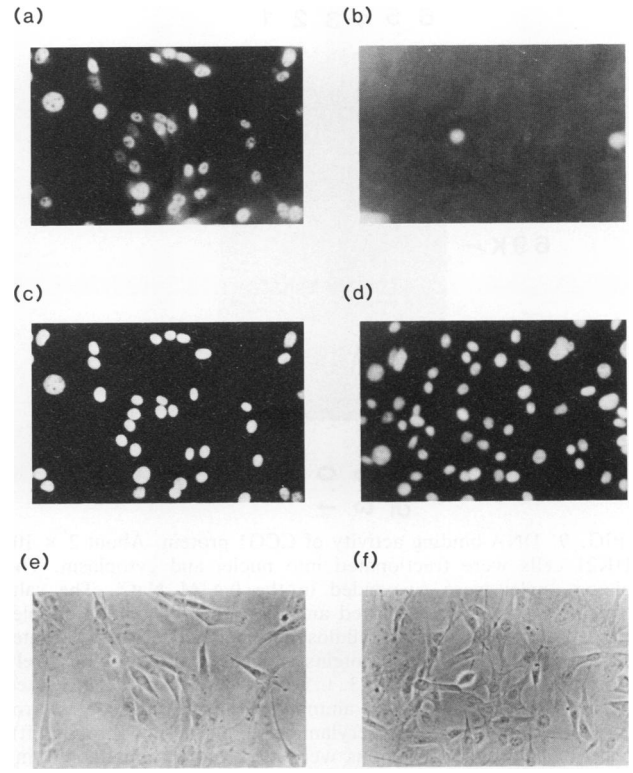

FIG. 7. Immunofluorescence staining of BHK cells with the anti-CCG1 antibody. BHK cells exponentially growing on the coverslip were fixed and treated sequentially by anti-CCG1 (a) or anti-CCG1 plus synthetic peptide as a competitor (b) and then by rhodamine-conjugated goat anti-rabbit immunoglobulin G. The same preparations were stained with Hoechst 33258 (c and d). A view of the fixed BHK21 cells obtained through phase-contrast microscopy is shown (e and f). Panels a, c, and e and b, d, and f show the same fields.

DNA column and was eluted in the presence of 0.3 M NaCl, a concentration similar to that used for the extraction from the nuclei (Fig. 8).

DISCUSSION

We previously cloned a CCG1 cDNA from a cDNA library of the secondary ts^+ transformants of *ts*BN462, K-1-1 cells. This cDNA library consists mainly of hamster cDNA. In such a cDNA library, the isolated human cDNA probably contained hamster sequences either at both ends or at one end, which were transcribed from the hamster DNA flanking the integrated human DNA. As predicted, the previously cloned CCG1 cDNA was found to contain hamster cDNA at both the 5′ and the 3′ ends.

Human DNA of 120 kb was isolated from a secondary ts^+ transformant of *ts*BN462, K-1-2. On the basis of the restriction map, the 5′ part of this human DNA appears to be identical with the previously cloned *CCG1* genomic DNA which was present in all ts^+ transformants of *ts*BN462. However, its 3′ 50 kb were present in only a series of ts^+ transformants, indicating that it is not essential to complement the *ts*BN462 mutation. Since the contiguous cDNA transcribed from both the 5′ and the 3′ parts of the newly isolated human DNA was cloned from a human cDNA library, it was concluded that the real human *CCG1* gene has a length of 120 kb. The newly isolated *CCG1* cDNA, *CCG1*p, was predicted to encode a protein of 210 kDa. Consistent with this, a protein with molecular mass slightly greater than 200 kDa was specifically identified in human, monkey, and hamster cells by immunoblotting.

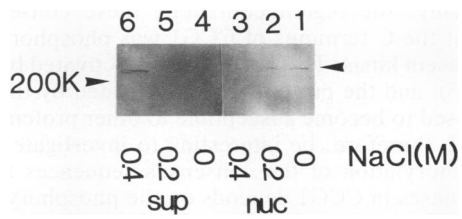

FIG. 8. Salt extraction of CCG1 from isolated nuclei of BHK21 cells. Nuclei (2×10^6) isolated from BHK21 cells were incubated in the presence of NaCl at concentrations ranging from 0.0 to 0.4 M for 10 min at 0°C and were then centrifuged. Proteins retained in the nuclei (lanes 1 through 3) and released into the supernatant (lanes 4 through 6) were analyzed by immunoblotting. The molar concentrations (M) of NaCl are shown below the lanes, and the position of p210 is shown by the arrowhead on the right.

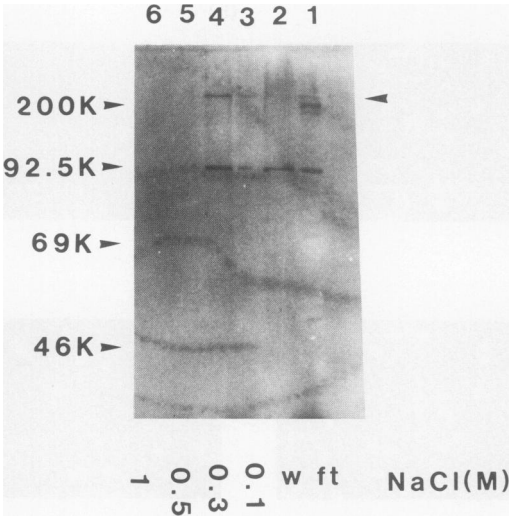

FIG. 9. DNA-binding activity of CCG1 protein. About 2×10^7 BHK21 cells were fractionated into nuclei and cytoplasm. The isolated nuclei were suspended in the 0.4 M NaCl. The salt-extracted proteins were diluted and then charged onto a double-stranded calf thymus DNA-cellulose column as described in Materials and Methods. Bound proteins were then eluted successively with buffer containing 0.1, 0.3, 0.5, and 1.0 M NaCl. The eluted proteins were precipitated by ammonium sulfate and then electrophoresed on 8% SDS-polyacrylamide gel. The flowthrough (ft), wash (w), and eluted fractions were analyzed by immunoblotting with the anti-CCG1 antibody. The molar concentrations (M) of NaCl are shown below the lanes, and the position of p210 is shown by the arrowhead on the right.

The 14-kb *Eco*RI fragment of the *CCG1* gene is absent in a series of ts^+ transformants of *ts*BN462. Conversely, the 17-kb fragment of the *Eco*RI fragment lying to the 5′ side of the 14-kb *Eco*RI fragment (Fig. 1) was retained in all ts^+ transformants through DNA-mediated gene transfer. Thus, the *CCG1* gene seems to have a tendency to be disrupted between the 17- and 14-kb *Eco*RI fragment (Fig. 1), and the 3′ region of the *CCG1* gene from the 14-kb *Eco*RI fragment downstream is not required to complement the *ts*BN462 mutation. The newly identified 3′ region of the *CCG1* gene encodes an acidic peptide containing four glutamic or aspartic acid clusters flanked with serine. This is a typical target of casein kinase II, which phosphorylates topoisomerase II and RNA polymerase (1, 7) in addition to several proteins implicated in neoplastic transformation, such as the Myc, Myb, and the E7 of the human papilloma virus type 16 (2, 17, 18). Actually, the region containing these consensus sequences at the C terminus of CCG1 was phosphorylated in vitro by casein kinase II. This enzyme is activated by growth factors (35), and the proteins phosphorylated by this kinase are supposed to become susceptible to other protein kinases (9). It will, therefore, be interesting to investigate whether the phosphorylation of the consensus sequences for other protein kinases in CCG1 depends on the phosphorylation of the casein kinase II target sequences. Since casein kinase II may be involved in the early regulation of the cell cycle, the newly cloned C-terminal region of the CCG1 gene may be a regulatory domain of CCG1.

On the basis of the putative amino acid sequence, CCG1 was found to have several of the characteristic structures of nuclear proteins, such as a proline cluster, a nuclear location signal, and a sequence similar to the DNA-binding domain of HMG1. In fact, CCG1 is actually located in the nucleus and has DNA-binding activity.

Proline clusters have been found in site-specific transcription factors such as the CTF/NF-1 family protein (20) and in enhancer-binding factor such as AP-2 (37), suggesting that CCG1 may be involved in transcription. The overall structure of CCG1 is roughly similar to those of HMG1 and hUBF, which have internal HMG1 boxes and acidic tails (14). Thus, CCG1 may be a new transcription factor needed in the process of G_1 phase.

ACKNOWLEDGMENTS

We thank H. Okazaki (Hoechst Japan) for preparing the synthetic peptide. We also thank Y. Taya (National Cancer Institute) for suggestions about the casein kinase II. We also thank K. Miller for proofreading the English in the manuscript.

This work was supported by a grant-in-aid for Scientific and Cancer Research from the Ministry of Education, Science and Culture of Japan, by the Science and Technology Agency, and by the Mitsubishi Foundation.

REFERENCES

1. **Ackerman, P., C. V. C. Glover, and N. Oscharoff.** 1988. Phosphorylation of DNA topoisomerase II in vivo and in total homogenates of Drosophila KC cells. J. Biol. Chem. **263:**12653–12660.
2. **Barbosa, M. S., C. Edmonds, C. Fisher, J. T. Schiller, D. R. Lowy, and K. H. Vousden.** 1990. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and SV40 large T antigen contains separate domains for RB binding and casein kinase II phosphorylation. EMBO J. **9:**153–160.
3. **Basilico, C.** 1977. Temperature-sensitive mutations in animal cells. Adv. Cancer Res. **24:**223–266.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
5. **Brown, C. J., T. Sekiguchi, T. Nishimoto, and H. F. Willard.** 1989. Regional localization of CCG1 gene which complements hamster cell cycle mutation BN462 to Xq11-Xq13. Somatic Cell Genet. **15:**93–96.
6. **Burnette, W. N.** 1981. "Western Blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal. Biochem. **112:**195–203.
7. **Dahmus, M. E.** 1981. Phosphorylation of eukaryotic DNA-dependent RNA polymerase. J. Biol. Chem. **256:**3332–3339.
8. **Edelman, A. M., D. A. Blumenthal, and E. G. Krebs.** 1987. Protein serine and threonine kinases. Annu. Rev. Biochem. **56:**563–613.
9. **Fiol, C. J., A. M. Mahreholz, Y. Wang, R. W. Roeske, and P. J. Roach.** 1987. Formation of protein kinase recognition site by covalent modification of the substrate. J. Biol. Chem. **262:**14042–14048.
10. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85:**8998–9002.
11. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23:**175–182.
12. **Hartwell, L. H., J. Culotti, J. Pringle, and B. Reid.** 1974. Genetic control of the cell division cycle in yeast. Science **183:**46–51.
13. **Helmann, J. D., and M. J. Chamberlin.** 1988. Structure and function of bacterial sigma factors. Annu. Rev. Biochem. **57:**839–872.
14. **Jantzen, H.-M., A. Admon, S. P. Bell, and R. Tjian.** 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature (London) **344:**830–836.
15. **Kalderon, D., B. Roberts, W. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear

location. Cell **39**:499–509.
16. **Liu, H. T., C. W. Gibson, R. R. Hirschhorn, S. Rittling, R. Baserga, and W. E. Mercer.** 1985. Expression of thymidine kinase and dihydrofolate reductase genes in mammalian ts mutants of the cell cycle. J. Biol. Chem. **260**:3269–3274.
17. **Luescher, B., E. Christenson, D. W. Litchfield, E. G. Krebs, and R. N. Eisenman.** 1990. Myb DNA binding inhibited by phosphorylation at a site deleted during oncogenic activation. Nature (London) **344**:517–522.
18. **Luescher, B., E. A. Kuenzel, E. G. Krebs, and R. N. Eisenman.** 1989. Myc oncoproteins are phosphorylated by casein kinase II. EMBO J. **8**:1111–1119.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58**:741–753.
21. **Nishimoto, T., T. Sekiguchi, R. Kai, K. Yamashita, T. Takahashi, and M. Sekiguchi.** 1982. Large-scale selection and analysis of temperature-sensitive mutants for cell reproduction from BHK cells. Somatic Cell Genet. **8**:811–824.
22. **Nurse, P., P. Thuriaux, and K. Nasmyth.** 1976. Genetic control of the division cycle of the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. **146**:167–178.
23. **Ohtsubo, M., R. Kai, N. Furuno, T. Sekiguchi, M. Sekiguchi, H. Hayashida, K. Kuma, T. Miyata, S. Fukushige, T. Murotsu, K. Matsubara, and T. Nishimoto.** 1987. Isolation and characterization of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation. Genes Dev. **1**:585–593.
24. **Ohtsubo, M., H. Okazaki, and T. Nishimoto.** 1988. The RCC1 protein, a regulator for the onset of chromosome condensation locates in the nucleus and binds to DNA. J. Cell Biol. **109**:1389–1397.
25. **Prescott, D. M.** 1976. The cell cycle and the control of cellular reproduction. Adv. Genet. **18**:91–177.
26. **Randall, K., R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
27. **Ruther, U., and B. Muller-Hill.** 1983. Easy identification of cDNA clones. EMBO J. **2**:1791–1794.
28. **Sanger, F. S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
29. **Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian.** 1988. A family of human CCAAT-box binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) **334**:218–224.
30. **Schluter, S. M., and J. J. Marchalonis.** 1986. Antibodies to synthetic joining segment peptide of the T-cell receptor β-chain: serological cross-reaction between products of T-cell receptor genes, antigen binding T-cell receptors, and immunoglobulins. Proc. Natl. Acad. Sci. USA **83**:1872–1876.
31. **Sekiguchi, T., T. Miyata, and T. Nishimoto.** 1988. Molecular cloning of the cDNA of human X chromosomal gene (CCG1) which complements the temperature-sensitive G1 mutants, tsBN462 and ts13, of the BHK cell line. EMBO J. **7**:1683–1687.
32. **Sekiguchi, T., M. Yoshida, M. Sekiguchi, and T. Nishimoto.** 1987. Isolation of a human X chromosome-linked gene essential for progression from G1 to S phase of the cell cycle. Exp. Cell Res. **169**:395–407.
33. **Shinkawa, K., S. Nakajo, K. Nakaya, and Y. Nakamura.** 1987. Purification of substrate proteins of casein kinases from the cytosol fraction of AH-66 hepatoma cells. Biochim. Biophys. Acta **930**:446–453.
34. **Sommercorn, J., J. A. Mulligan, F. J. Lozeman, and E. G. Krebs.** 1987. Activation of casein kinase II in response to insulin and to epidermal growth factor. Proc. Natl. Acad. Sci. USA **84**:8834–8838.
35. **Suzuki, M.** 1988. SPXX, a frequent sequence motif in gene regulatory proteins. J. Mol. Biol. **207**:61–84.
36. **Th'ng, J., P. Wright, J. Hamaguchi, M. G. Lee, C. J. Norbury, P. Nurse, and E. M. Bradbury.** 1990. The FT210 cell line is a mouse G2 phase mutant with a temperature-sensitive CDC2 gene product. Cell **63**:313–324.
37. **Williams, T., A. Adamon, B. Luescher, and R. Tjian.** 1988. Cloning and expression of AP-2, a cell type-specific transcription factor that activates inducible enhancer elements. Genes Dev. **2**:1557–1569.