1700–1706 *Nucleic Acids Research, 1998, Vol. 26, No. 7*

Isolation and characterization of a human cDNA for mRNA 5′-capping enzyme

Toshiko Yamada-Okabe, Rikuo Doi, Osamu Shimmi[1], Mikio Arisawa[1] and Hisafumi Yamada-Okabe[1,*]

Department of Hygiene, School of Medicine, Yokohama City University, 3-9, Fukuura, Kanazawa, Yokohama 236, Japan and [1]Department of Mycology, Nippon Roche Research Center, 200 Kajiwara, Kamakura, Kanagawa 247, Japan

Received December 5, 1997; Revised and Accepted February 16, 1998 DDBJ/EMBL/GenBank accession nos AB009022–AB009024

ABSTRACT

The human mRNA 5′-capping enzyme cDNA was identified. Three highly related cDNAs, *HCE1* (human mRNA capping enzyme 1), *HCE1A* and *HCE1B*, were isolated from a HeLa cDNA library. The *HCE1* cDNA has the longest ORF, which can encode a 69 kDa protein. A short region of 69 bp in the 3′-half of the *HCE1* ORF was missing in *HCE1A* and *HCE1B*, and, additionally, *HCE1B* has an early translation termination signal, which suggests that the latter two cDNAs represent alternatively spliced product. When expressed in *Escherichia coli* as a fusion protein with glutathione *S*-transferase, Hce1p displayed both mRNA 5′-triphosphatase (TPase) and mRNA 5′-guanylyltransferase (GTase) activities, and it formed a cap structure at the 5′-triphosphate end of RNA, demonstrating that it indeed specifies an active mRNA 5′-capping enzyme. The recombinant proteins derived from *HCE1A* and *HCE1B* possessed only TPase activity. When expressed from *ADH1* promoter, *HCE1* but not *HCE1A* and *HCE1B* complemented *Saccharomyces cerevisiae CEG1* and *CET1*, the genes for GTase and TPase, respectively. These results demonstrate that the N-terminal part of Hce1p is responsible for TPase activity and the C-terminal part is essential for GTase activity. In addition, the human TPase domain cannot functionally substitute for the yeast enzyme *in vivo*.

INTRODUCTION

Almost all eukaryotic mRNAs are capped at their 5′-termini (1). Such modification by capping has been shown to be crucial for the stabilization (2–4), processing (5–10), nuclear export (11), and translation of mRNA (see refs 12,13 for reviews). Capping occurs at an early stage of transcription and requires at least three enzymatic activities, mRNA 5′-triphosphatase (termed TPase in this study), mRNA 5′-guanylyltransferase (termed GTase in this study), and cap methyltransferase (termed MTase in this study) (see refs 14,15 for reviews). GTase and TPase activities are associated in one enzyme called mRNA 5′-capping enzyme (termed capping enzyme in this study), but MTase activity is dissociated from these two activities (16,17).

In *Saccharomyces cerevisiae* the capping enzyme consists of two subunits, α and β. The 52 kDa α-subunit is authentic for GTase activity, and the 80 kDa β-subunit is responsible for TPase activity (18). Three yeast GTase genes have been cloned from *S.cerevisiae* (19), *Schizosaccharomyces pombe* (20) and *Candida albicans* (21), and their products were found to possess only GTase activity, demonstrating that subunit structure is a characteristic of yeast capping enzymes. Recently *CET1* was identified as the gene for the *S.cerevisiae* TPase (β subunit of capping enzyme) (22).

In addition to yeast GTases, chlorella virus PBCV-1 GTase was identified as a 38 kDa protein (23). This is the smallest protein reported as a GTase to date, and no TPase activity is associated with this protein. The finding of PBCV-1 GTase implies that the chlorella virus capping enzyme has a subunit structure similar to those of the yeast capping enzymes, and the crystal structure of chlorella virus PBCV GTase is solved (24). The enzyme consists of two domains with a cleft between them. The structure of GTase resembles that of T7 DNA ligase; T7 DNA ligase also comprises two domains which produce a cleft between them (25). The cleft observed in GTase, however, is much deeper and narrower than that of DNA ligase, which may account for the substrate specificity of capping enzyme for single strand RNA molecules. After binding to manganese ions, the enzyme undergoes a big conformational change from open state to closed state, which may facilitate the hydrolysis of β and γ phosphoryl groups of bound GTP to form a stable enzyme–GMP complex (24).

On the other hand, mammalian capping enzyme was purified as a 69 kDa single polypeptide having both TPase and GTase activities (26). Recently, a cDNA for the *CEL1* gene was isolated from *Caenorhabiditis elegans* (27). In the putative protein Cel1p, the C-terminal part is highly related to yeast and viral GTases, while the N-terminal region contains an amino acid sequence motif that is characteristic of the active site of protein tyrosine phosphatase families. Despite the sequence similarity of Cel1p to the protein tyrosine phosphatases, a recombinant Cel1p does not recognize phosphotyrosine, but instead catalyzes the hydrolysis of a γ-phosphoryl group at the 5′-termini of RNAs, suggesting

*To whom correspondence should be addressed. Tel: +81 467 47 2213; Fax: +81 467 46 5320; Email: hisafumi.okabe@roche.com

that protein tyrosine phosphatases and TPase share a similar catalytic mechanism (27). Although GTase activity of Cel1p remains to be established, all of these facts support the idea that Cel1p is a capping enzyme of *C.elegans*.

To gain more insight into the characteristics of a higher eukaryotic capping enzyme, we have isolated a cDNA for the human capping enzyme. Three highly related cDNAs were identified from a HeLa cDNA library. Comparison of the amino acid sequences of the expected products of these cDNAs revealed that two of them have deletions in the C-terminal half of the longest cDNA product, and that only the longest cDNA product has GTase activity, whereas all three cDNA products are active as TPase. Moreover, the human capping enzyme cDNAs defective in GTase did not functionally complement the *S.cerevisiae CET1*. Thus, it seems that the human capping enzyme mRNA is alternatively spliced producing GTase-deficient enzymes, and that the human and yeast TPases are functionally diverged.

MATERIALS AND METHODS

Screening of a HeLa cDNA library

The KXDG motif characteristic of proteins forming a stable enzyme–nucleotide monophosphate complex was used to search human genome and cDNA databases, and a short cDNA sequence in the Incyte Lifeseq™ database (http://www.incyte.com/) was found to harbor a sequence of ADG. Using the nucleotide sequence of this clone, a 0.5 kb DNA fragment was amplified from a HeLa cDNA library by PCR, and it was used as a probe for cloning a full length cDNA. Primers used to amplify the probe DNA were, 5′-CCCGGGAATTCAGACTGGGAGGCTTCCATT-3′ and 5′-CCCGGGAATTCGGCTCGAGAGCAGATGGTA-3′. Hybridization and washing of the filters were carried out under stringent conditions [20 mM sodium phosphate (pH 7.2), 5× SSC (1× SSC contains 150 mM NaCl and 15 mM sodium citrate), 5× Denhardt's solution, 0.1% (w/v) SDS, 50% (v/v) formamide at 42°C for hybridization; 0.1× SSC and 0.1% SDS at 50°C for washing], and phages that were strongly hybridized with the probe DNA were collected. After a third screening, phages from the positive plaques were expanded, and the phage DNA was extracted (28). Then cDNA insert was excised from the λgt10 vector by digesting the phage DNA with *Not*I and *Sal*I, and cloned at the *Sma*I site of pUC19. Radiolabelling the probe DNA was performed by the random priming method using [α-^{32}P]dCTP (28), and DNA sequencing was carried out as described elsewhere (28). A HeLa cDNA library was purchased from Clontech (USA).

Expression and purification of the recombinant capping enzyme

For subcloning the human capping enzyme cDNAs, *Eco*RV sites were generated at both ends of ORFs using *Eco*RV linkers. The coding regions of the human capping enzyme cDNAs (*HCE1*, *HCE1A* and *HCE1B*) and the *S.cerevisiae CEG1* gene were cloned at *Bam*HI and *Sma*I sites of pGEX2T, respectively (29), and the resulting plasmids were transfected into *Escherichia coli* JM109 to let them express recombinant yeast and human capping enzymes as a fusion with GST. Induction of expression of recombinant capping enzymes were carried out with IPTG as described (29). At 4 h after the addition of IPTG, the bacterial cells were harvested, suspended in buffer A containing 20 mM Tris–HCl (pH 7.5), 0.5 mM EDTA, 50 mM NaCl, 10 mM β-mercaptoethanol, 10%(v/v) glycerol, 0.05% (v/v) NP-40 and 1 mM PMSF (phenylmethylsulfonyl fluoride), and lysed by sonication. Cell debris were removed by centrifugation at 15 000 *g* at 4°C for 30 min, and the supernatant was added to glutathione Sepharose CL-4B beads that were equilibrated with buffer A. After a gentle rotation at 4°C for 1 h, the beads were extensively washed with buffer A and subjected to western blotting or enzyme assays.

Assays for capping enzyme

An assay for enzyme–GMP complex formation was performed in a standard reaction mixture containing 20 mM Tris–HCl (pH 7.5), 3 mM $MgCl_2$, 10 mM dithiothreitol, 0.4 mg/ml bovine serum albumin, 2 mg/ml inorganic pyrophosphatase, 20%(v/v) glycerol, 1 mM PMSF, 1 mM [α-^{32}P]GTP (specific activity 1 × 10^5 c.p.m./pmol), and ~0.1 μg protein of the fusion proteins that bound to the glutathione Sepharose CL-4B beads at 30°C for 10 min (30). After the reaction, the proteins were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), and the proteins that were covalently bound to GMP were visualized by autoradiography.

For a TPase assay, the glutathione Sepharose CL-4B beads in complex with ~0.1 μg protein of the purified fusion proteins were washed three times with buffer B containing 50 mM Tris–HCl (pH 7.9) and 2 mM dithiothreitol and were then incubated with [γ-^{32}P]ATP-terminated RNA at 30°C for 30 min (26). After removing the beads by a brief centrifugation, the reaction mixture was analysed by TLC using polyethyleneimine cellulose plates, and the spots were visualized by Fuji BAS 2000. The [γ-^{32}P]ATP-terminated RNA was prepared as follows. Heat-denatured calf thymus DNA (50 μg) was incubated in a buffer containing 100 mM Tris–HCl (pH 7.9), 2 mM dithiothreitol, 5 mM $Mg(OAc)_2$, 2 mM $MnCl_2$, 1 mM [γ-^{32}P]ATP (1000–5000 c.p.m./pmol), and 1 U *E.coli* RNA polymerase at 37°C for 90 min (26). After extracting with phenol/chloroform, [γ-^{32}P]ATP-terminated RNA was precipitated with ethanol and separated from [γ-^{32}P]ATP by Sephadex G-25 column chromatography.

For the *in vitro* cap formation assay, the glutathione Sepharose CL-4B beads complexed with ~0.1 μg protein of the fusion proteins were incubated in a buffer containing 10 mM Tris–HCl (pH 7.5), 3 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 20%(v/v) glycerol, 2.4 mM [α-^{32}P]GTP (400 Ci/mmol), and ~50 pmol pppA-terminated RNA at 30°C for 60 min. After reaction, the RNA was digested with nuclease P1 and calf intestine alkaline phosphatase as described (16). The resulting nucleotides were separated on DE81 paper by paper electrophoresis and visualized by Fuji BAS 2000.

Western blotting

The purified recombinant capping enzymes bound to the glutathione Sepharose CL-4B beads were suspended in a loading buffer for SDS–PAGE and heated at 90°C for 2 min. After the beads were removed by a brief centrifugation, the proteins were fractionated by SDS–PAGE, transferred to a polyvinilidene difluoride (PVDF) membrane electrophoretically (28), and reacted with anti-GST antibody (Pharmacia) and then with horseradish-peroxidase conjugated protein A (Amersham) as described in the manual provided by the supplier. The protein hybridized with anti-GST-antibody was detected by an ECL-protein detection kit (Amersham).

Figure 1. Structure and amino acid sequence of human capping enzyme. (**A**) Structures of Hce1p, Hce1ap and Hce1bp are shown. The positions of the internal deletion and the early translation termination found in Hce1ap and Hce1bp are indicated. A proline rich sequence at the boundary of the TPase and GTase domains and the KXDG motif in the GTase domain of hce1p are also shown. Arrows at the bottom represent the positions and directions of the primers used for RT-PCR. (**B**) Amino acid sequence of Hce1p is compared with those of a probable *C.elegans* capping enzyme (Cel1p), *S.pombe* GTase (Pce1p), *S.cerevisiae* GTase (Ceg1p) and *C.albicans* GTase (Cgt1p). Identical and homologous amino acids among these proteins are marked by asterisks and dots, respectively. Domains that are highly conserved in all the proteins are numbered from I to IX according to their order from the N-terminus. Amino acid sequences are aligned using the BLAST and FAST programs.

Site-directed mutagenesis

The mutant human capping enzymes bearing alanine substitution for K294, R299, E345, K458 or K460 were generated by the oligonucleotide-mediated mutagenesis of the human capping enzyme cDNA with Mutan-Express Km™ (Takara) (28). The entire ORF of the human capping enzyme cDNA was cloned at the *Sal*I site of pKF18k (Takara), and hybridized with an appropriate primer. The resulting mutant human capping enzyme cDNAs were tested for their ability to complement the *S.cerevisiae CEG1* gene.

Reverse transcribed polymerase chain reaction (RT-PCR)

RT-PCR was carried out to quantify the human capping enzyme mRNA. Total RNA was prepared from HeLa cells by the guanidium-isothiocyanate method and applied to an oligo(dT) cellulose column (28). Single-stranded complementary DNA was made from poly(A)$^+$ RNA using AMV reverse transcriptase and oligo(dT) primer followed by the indicated cycles of the incubation at 94°C for 1 min, at 55°C for 1 min, and at 72°C for 2 min. The primers used for PCR were: 5′-GTGATTTTAATGTTC-GTCTGCAGTGTATAG-3′ for primer a; 5′-GCCAAAGAAGT-GAGCCATGAAATGGATGGA-3′ for primer b; 5′-GACGAAA-ATCCACAGAATTCAGACTGGGAG-3′ for primer c; and 5′-CT-TTTGTCTGTTCTCTGTCTCATGAAGACC-3′ for primer d.

Generation of yeast strains

The ORFs of the *S.cerevisiae CEG1* and *CET1* genes, and the human capping enzyme cDNAs were cloned between *Hin*dIII and *Pst*I site (between the *ADH1* promoter and terminator) of pGBT9 that carried *TRP1* (Clontech), and the resulting plasmids were transfected into the *ceg1*Δ (*MATa ura3 leu2 his3 trp1 ceg1::LEU2 CEG1-URA3*) (21) and *cet1*Δ (*MATa ade2 ura3 leu2 his3 trp1 cet1::LEU2 CET1-URA3*) null mutant strains, in which the endogenous *CEG1* or *CET1* was disrupted by *LEU2*, but where episomal copies of *CEG1* or *CET1* cloned in YEp24 were maintained, respectively. Thus, the resulting *ceg1*Δ and *cet1*Δ null mutant strains grew in the absence of 5-fluoroorotic acid (5-FOA) but not in the presence of 5-FOA. After selecting trp$^+$ transformants, the ability of the human capping enzyme cDNAs to complement *CEG1* and *CET1* was examined by culturing them on an agar plate containing 5-FOA.

RESULTS

Cloning of the human capping enzyme cDNA

Enzymes such as capping enzyme and DNA ligase form a stable complex with nucleotide monosphosphate as a reaction intermediate, and the KXDG motif, where X stands for any amino acid, is known to be essential for the formation of the enzyme–nucleotide monophosphate complex (20,31–33). In an attempt to identify a human capping enzyme cDNA, we searched for a protein harboring the KXDG motif in human genomic and cDNA databases and found that a short cDNA fragment in the Incyte cDNA database could encode a protein starting with ADG. As this clone contained a DNA fragment of 0.5 kb and an ORF could extend in both 5′- and 3′-directions, it should be a partial cDNA clone. To clone a full length cDNA, a HeLa cDNA library was screened with the cDNA fragment that was amplified by PCR using the above sequence, and three different but highly related cDNA clones were isolated. We designated these cDNAs as *HCE1* (human mRNA capping enzyme 1), *HCE1A* and *HCE1B*. Sequencing of theses cDNAs revealed that *HCE1* contained the longest ORF, which can specify a 69 kDa protein. In both *HCE1A* and *HCE1B*, the region between positions 1270 and 1338 of the *HCE1* ORF was missing, which eliminates 23 amino acids from the C-terminal half of the *HCE1* product (Hce1p) (Fig. 1A). In addition to the above internal deletion, *HCE1B* has a translation termination codon corresponding to position 1441 of the *HCE1* ORF, which deletes 117 C-terminal amino acids from Hce1p (Fig. 1A). Although nucleotide sequences of the 3′-untranslated regions of *HCE1* and *HCE1A* are identical, that of *HCE1B* is completely different from the other two clones. Otherwise, all these three cDNAs possess identical nucleotide sequences, strongly suggesting that they are alternatively spliced products.

HCE1 is the most probable candidate for the human capping enzyme cDNA because (i) the predicted molecular weight of Hce1p coincides with that of the partially purified human capping enzyme (26), (ii) Hce1p is highly related to the probable *C.elegans* capping enzyme Cel1p (27) over the entire protein, (iii) the C-terminal part of Hce1p shares high degrees of sequence identity with yeast and viral GTases (19–21,34–36), and (iv) the KXDG motif characteristic of the enzyme–GMP complex formation is conserved in Hce1p (20,31–33) (Fig. 1B).

Next, we intended to establish the enzyme activities of Hce1p. Hce1p, Hce1ap, Hce1bp and the *S.cerevisiae* GTase (Ceg1p) were expressed in bacterial cells as a fusion with glutathione *S*-transferase (GST), and the proteins were purified by glutathione Sepharose beads. Western blotting with anti-GST antibody confirmed the similar levels of each protein, but only Ceg1p and Hce1p formed a stable enzyme–GMP complex when incubated with [α-^{32}P]GTP (Fig. 2). This result indicates that Hce1p harbors GTase activity and that the short internal deletion occurring in *HCE1A* and *HCE1B* abolished the GTase activity of Hce1p.

The N-terminal portion of Hce1p is highly homologous to that of Cel1p. Because Cel1p is shown to possess TPase activity (27), we also looked for the TPase activity of Hce1p. As expected, Hce1p, Hce1ap and Hce1bp all efficiently removed the ^{32}P-labelled γ-phosphate at the 5′-terminus of RNA, while GST and Ceg1p did not. To further confirm the TPase activity of these proteins, a cap formation assay was performed using an RNA substrate with 5′-triphosphate ends. This assay should allow us to exclude the possibility of general phosphatase activity of these proteins,

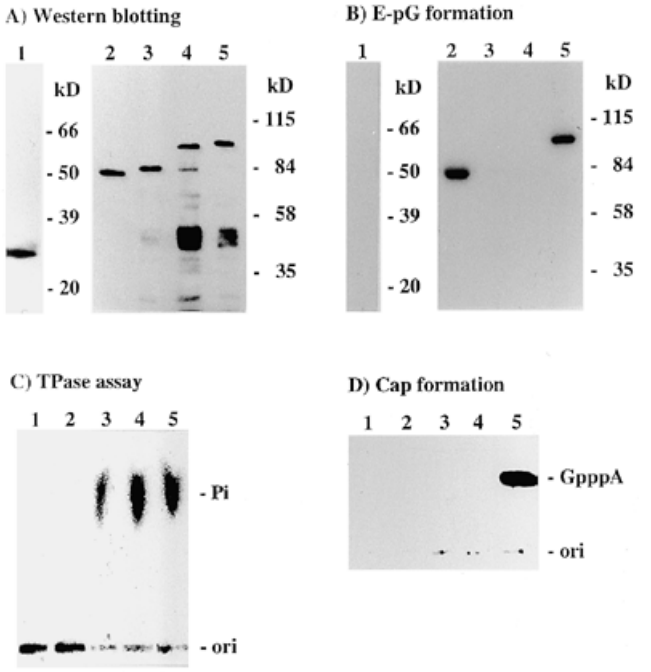

Figure 2. Expression and enzyme activities of the human capping enzyme cDNAs. Ceg1p, Hce1p, Hce1ap and Hce1bp were expressed in *E.coli* as a fusion with GST and purified with glutathione Sepharose beads. Approximately 0.1 µg of the purified proteins bound to the beads were detected by western blotting using anti-GST antibody (**A**), or subjected to the assays of enzyme–GMP complex (E-pG) formation (**B**) TPase (**C**) and cap formation at the 5′-triphosphate ends of RNA (**D**). For western blotting and E-pG formation, the proteins were fractionated on 12.5% (for GST) or 10% (for other proteins) SDS–polyacrylamide gels. Lane 1, GST; lane 2, GST-Ceg1p; lane 3, GST-Hce1ap; lane 4, GST-Hce1bp; lane 5, GST-Hce1p.

because guanylylation at the 5′-end of RNA requires diphosphate-ends of RNA. As shown in Figure 2, Hce1p formed a cap structure at the 5′-triphosphate ends of an RNA substrate, demonstrating that *HCE1* indeed encodes a human capping enzyme. Failure of Hce1ap and Hce1bp to cap the 5′-triphosphate ends of RNA is due to their lack of GTase activity. Taken together, we concluded that *HCE1* is a cDNA for human capping enzyme.

Complementation of yeast *CEG1* by *HCE1*

Comparison of the amino acid sequences of Hce1p, *C.elegans* Cel1p, and the yeast GTases revealed nine domains that are quite conserved among the yeast and higher eukaryotic GTases (Fig. 1B). For convenience, these nine domains are referred to as domains I–IX according to their order from the N-terminus. Domain II contains the KXDG motif in which lysine at 294 is thought to form a phosphoamide bond with GMP (32). According to the structure of PBCV-1 GTase, arginine at 299 in domain II and glutamic acid at 345 in domain III may form hydrogen bonds with hydroxyl groups of ribose of GTP. Further, lysine at 458 and at 460 in domain VI may also interact with the α-phosphate of GTP in the closed form and in the open form, respectively (24).

Because all these important amino acids are conserved in Hce1p, we asked whether human *HCE1* functionally substitutes for the yeast *CEG1* gene. The *S.cerevisiae ceg1*Δ null mutant strain in which the endogenous *CEG1* gene was disrupted, but where episomal copies of *CEG1* were maintained in an

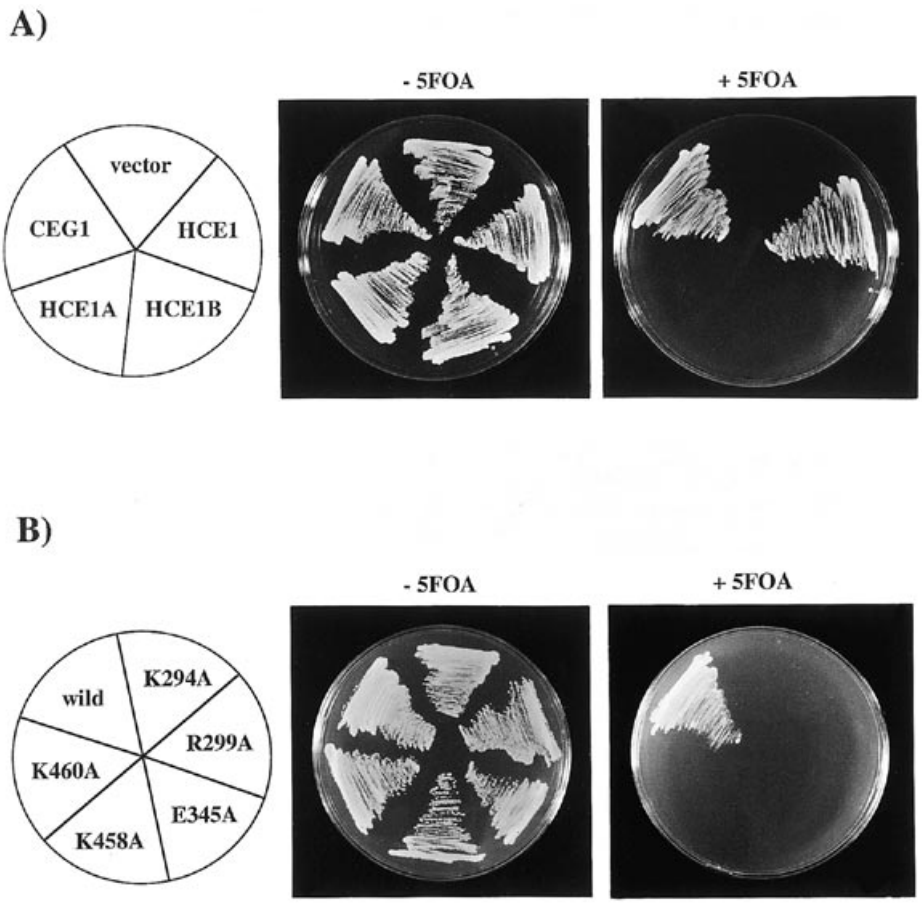

Figure 3. Functional complementation of *S.cerevisiae CEG1* by the human capping enzyme cDNA. (**A**) *Saccharomyces cerevisiae ceg1*Δ*::LEU2* cells that harbored *CEG1* in YEp24 were further transformed with pGBT9 bearing *CEG1*, *HCE1*, *HCE1A* or *HCE1B*. (**B**) The wild-type and mutant *HCE1* cDNAs bearing alanine substitution for K294 (K294A), R299 (R299A), E345 (E345A), K458 (K458A) or K460 (K460A) were cloned in pGBT9 and transfected into the *S.cerevisiae ceg1*Δ*::LEU2* cells. Transfectants were spread on agar plates with (+) or without (–) 5-FOA, and incubated at 30°C for 3 days.

URA3-containing plasmid (YEp24), grew in the absence of 5-FOA, but died in the presence of 5-FOA from the lack of a functional *CEG1* gene (Fig. 3A) (21). Expression of *HCE1* under the control of *ADH1* promoter supported the growth of the *ceg1*Δ null mutant strain, even in the presence of 5-FOA, demonstrating that human capping enzyme functionally complements yeast GTase. *HCE1A* and *HCE1B* failed to rescue the *ceg1*Δ null mutants in the same experiment. Furthermore, none of the mutant *HCE1* genes encoding the protein with alanine substitution for K294, R299, E345, K458 or K460 complemented *CEG1* (Fig. 3B), confirming that all these conserved amino acids are essential for the GTase activity. These results are consistent with the previous findings that the proteins specified by *HCE1A* and *HCE1B* are defective in GTase activity and with the results of the mutation study with *CEG1* by Wang *et al*. (37).

Levels of the human capping enzyme mRNAs

As there are three mRNA species of human capping enzyme, we examined the levels of these mRNAs. Northern blotting of the human capping enzyme mRNA did not allow us to distinguish among the *HCE1*, *HCE1A* and *HCE1B* mRNAs. Therefore, we used RT-PCR with different sets of primers specific and not-specific to *HCE1* (Fig. 4). The combination of primers specific to *HCE1* (primers b and c) gave rise to only one band which fits the size estimated from the cDNA sequence (124 bp). The amount of PCR product increased in a PCR-cycle dependent manner; a weak band appeared after 25 cycles, and it became significantly thicker after 30 cycles. Thus, this system can be applied to estimate the relative abundance of the three mRNAs for human capping enzyme.

Primers for common sequences among the three cDNAs (primers a and c) produced two bands of 288 and 219 bp. From the cDNA sequences, the thick 288 bp band was amplified from *HCE1*, and the faint 219 bp band was formed from *HCE1A* and *HCE1B*. Use of the primers that hybridized with *HCE1* and *HCE1A* (primers a and d) also led to the amplification of two bands, the thick 478 bp band generated from *HCE1* and the faint 409 bp band produced from *HCE1A*. In both cases, the amounts of the DNA fragment amplified from *HCE1* were significantly higher than those from *HCE1A* and *HCE1B*. All these results indicate that mRNA for *HCE1* is predominantly expressed among the three human capping enzyme mRNA species.

Complementation of yeast *CET1* by *HCE1*

CET1 was recently identified as the *S.cerevisiae* TPase (β subunit of capping enzyme) gene (22). Cet1p, however, shares no

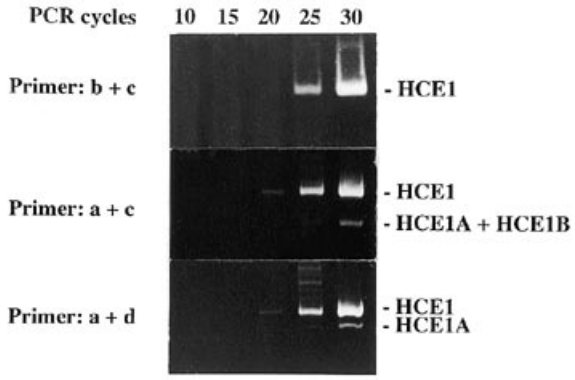

Figure 4. Comparison by RT-PCR of the *HCE1*, *HCE1A* and *HCE1B* mRNA levels. One microgram of poly(A)$^+$ RNA extracted from HeLa suspension cultures was used to synthesize single strand cDNA followed by amplification by PCR with the indicated combinations of the primers. After the indicated cycles, the amplified DNA fragments were separated on 5% polyacrylamide gels and visualized by ethidium bromide staining. Positions of the DNA fragments amplified from each cDNA are indicated. Positions and directions of the primers are shown in Figure 1.

apparent sequence homology with Hce1p and Cel1p, and has no known phosphatase motif. As mentioned before, both Hce1p and Cel1p contain a motif in the N-terminal TPase domains of amino acid sequence characteristic of the active site of tyrosine phosphatases, and the active site cysteine of the protein tyrosine phosphatases is also conserved in these domains (39). This prompted us to examine the ability of *HCE1* to functionally complement *CET1*. The cells of *cet1*Δ bearing the episomal copies of *CET1* or *HCE1* survived in the presence of 5-FOA, but those with *HCE1A* or *HCE1B* died in the same culture condition. All the cells grew well when 5-FOA was absent (Fig. 5). These results indicate that at least in the level induced by *ADH1* promoter, the TPase domain of the human capping enzyme cannot functionally substitute for the yeast TPase *in vivo*.

DISCUSSION

We have identified *HCE1*, a human cDNA for the mRNA-capping enzyme. Bacterially expressed Hce1p possessed both TPase and GTase activities, and it formed a cap structure at the 5′-triphosphate ends of RNA. This is the first study that demonstrates both TPase and GTase activities with a single type of cDNA. Human capping enzyme mRNA seems to be spliced differently to produce two additional mRNA species, which are represented by *HCE1A* and *HCE1B*. From the nucleotide sequences, the *HCE1A* and *HCE1B* products appear to have the same internal 23 amino acid deletion within the C-terminal part of Hce1p, and an additional 117 amino acids within the C-terminal of Hce1p are also missing in the *HCE1B* product. Both Hce1ap and Hce1bp displayed TPase activity, while neither of them had GTase activity. Thus demonstrating that, as in the *C.elegans* Cel1p, the N-terminal part of Hce1p is responsible for TPase activity, and the C-terminal part is required for the GTase activity (27). In addition, 23 amino acids between positions 1269 and 1338 were essential for GTase activity of Hce1p. This is also consistent with the results that *HCE1*, but not *HCE1A* and *HCE1B*, functionally complemented *S.cerevisiae CEG1*.

During preparation of this paper, Yue *et al*. and McCracken *et al*. reported the cDNAs for the human and mouse capping enzyme (37,40). The amino acid sequence of the reported human capping enzyme differs by only one amino acid from that of Hce1p. It is also demonstrated in the paper (37,40) that the C-terminal GTase domain is responsible for the binding to the phosphorylated C-terminal domain of the largest subunit of RNA polymerase II, which explains the specific capping of class II transcripts.

Cel1p is shown to have TPase activity (27), but whether GTase activity resides in Cel1p remains to be demonstrated. When the sequences of Cel1p and Hce1p are compared, 14 amino acids in the C-terminal of Hce1p are missing in Cel1p. Although the importance of the 14 C-terminal amino acids of Hce1p for GTase activity is not known, the absence of these amino acids in Cel1p may account for its lack of GTase activity. Furthermore, the *C.elegans* capping enzyme mRNA may also be alternatively spliced so as to produce the GTase-deficient but TPase positive enzymes.

CET1 turned out to be the gene for the *S.cerevisiae* TPase (22). In this paper, we found that among three human capping enzyme cDNAs only *HCE1* functionally complemented *CET1 in vivo*, while all the cDNA products were active as TPase *in vitro*. This was a rather unexpected result because Takagi *et al*. (27) demonstrated the whole capping enzyme activity *in vitro* by combining Ceg1p and Cel1p that is also defective in GTase activity. Because the amino acid sequence of Cet1p is diverged

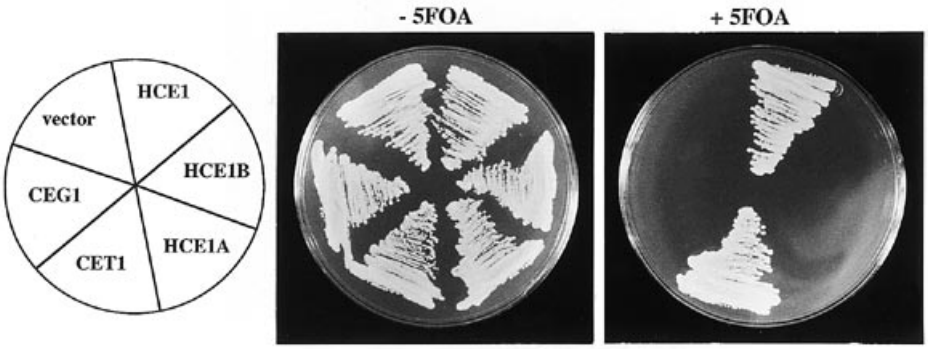

Figure 5. Functional complementation of *S.cerevisiae CET1* by the human capping enzyme cDNA. *Saccharomyces cerevisiae cet1*Δ*::LEU2* cells that harbored *CET1* in YEp24 were further transformed with pGBT9 bearing *CEG1*, *HCE1*, *HCE1A* or *HCE1B*. Transfectants were spread on agar plates with (+) or without (–) 5-FOA, and incubated at 30°C for 3 days.

from those of the higher eukaryotic enzymes, the catalytic mechanism of TPase may be different in yeast and higher eukaryotes, which may account for the failure of *HCE1A* and *HCE1B* to functionally complement *CET1 in vivo*. Another possibility may be that both α- and β-subunits should be associated with each other to be able to function as a capping enzyme *in vivo*, and that the human TPase domain has a low affinity for the yeast GTase.

The physiological roles of a GTase-deficient capping enzyme are unclear. From the experiments by RT-PCR, it appears that the level of mRNA for *HCE1* was significantly higher than those for *HCE1A* and *HCE1B*. This result implies that the enzyme possessing both TPase and GTase activities is predominantly expressed, and that the GTase-deficient enzymes specified by *HCE1A* and *HCE1B* are only minor populations. On the other hand, the molecular mass of the purified yeast capping enzyme is ~180 kDa (18). As the calculated molecular mass of α- and β-subunits of yeast enzyme are 52 and 62 kDa, respectively (19,22), the subunit structure of the yeast capping enzyme can be α1β2. Accordingly, human GTase-deficient enzymes may be associated with the native enzyme.

ACKNOWLEDGEMENTS

We thank S.Fukuchi for the database search and sequence alignment, and S.Miwa for reading the manuscript. This work was supported in part by a human science fund and a grant from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1 Shatkin,A.J. (1976) *Cell* **9**, 645–653.
2 Furuichi,Y., LaFiandra,A. and Shatkin,A.J. (1977) *Nature* **266**, 235–239.
3 Shimotohno,K., Kodama,Y., Hashimoto,J. and Miura,K.I. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2734–2738.
4 Murthy,K.G.K., Park,P. and Manley,J.L. (1991) *Nucleic Acids Res.* **19**, 2685–2692.
5 Edery,I. and Sonenberg,N. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7590–7594.
6 Inoue,K., Ohno,M., Sakamoto,H. and Shimura,Y. (1989) *Genes Dev.* **3**, 1472–1479.
7 Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell* **38**, 731–736.
8 Krainer,A.R., Maniatis,T., Ruskin,B. and Green,M.R. (1984) *Cell* **36**, 993–1005.
9 Patzelt,E., Thalmann,E., Hartmuth,K., Blaas,D. and Kuechler,E. (1987) *Nucleic Acids Res.* **15**, 1387–1399.
10 Ohno,M., Sakamoto,H. and Shimura,Y. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5187–5191.
11 Hamm,J. and Mattaj,I.W. (1990) *Cell* **63**, 109–118.
12 Filipowicz,W. (1978) *FEBS Lett.* **96**, 1–11.
13 Shatkin,A.J. (1985) *Cell* **40**, 223–224.
14 Mizumoto,K.and Kaziro,Y. (1987) *Prog. Nucleic Acids Res. Mol. Biol.* **34**, 1–28.
15 Shuman, S. (1995) *Prog. Nucleic Acids Res. Mol. Biol.* **50**, 101–129.
16 Mizumoto,K. and Lipmann,F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4961–4965.
17 Nishikawa,Y. and Chambon,P. (1982) *EMBO J.* **1**, 485–492.
18 Itoh,N., Yamada,H., Kaziro,Y. and Mizumoto,K. (1987) *J. Biol. Chem.* **262**, 1989–1995.
19 Shibagaki,Y., Itoh,N., Yamada,H., Nagata,S. and Mizumoto,K. (1992) *J. Biol. Chem.* **267**, 9521–9528.
20 Shuman,S., Liu,Y. and Schwer,B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12046–12050.
21 Yamada-Okabe,T., Shimmi,O., Doi,R., Mizumoto,K., Arisawa,M. and Yamada-Okabe,H. (1996) *Microbiology* **142**, 2515–2523.
22 Tsukamoto,T., Shibagaki,Y., Imajoh-Ohmi,S., Murakoshi,T., Suzuki,M., Nakamura,A., Gotoh,H. and Mizumoto,K. (1997) *Biochem. Biophys. Res. Comm.* **239**, 116–122.
23 Ho,C. K., van Etten,J.L. and Shuman,S. (1996) *J. Virol.* **70**, 6658–6664.
24 Hakansson,K., Doherty,A.J., Shuman,S. and Wigley,D.B. (1997) *Cell* **89**, 545–553.
25 Subramanya,H.S., Doherty,A.J., Ashford,S.R. and Wigley,D.B. (1996) *Cell* **85**, 607–615.
26 Yagi,Y., Mizumoto,K. and Kaziro,Y. (1983) *EMBO J.* **2**, 611–615.
27 Takagi,T., Moore,C.R., Diehn,F. and Buratowski,S. (1997) *Cell* **89**, 867–873.
28 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29 Smith,D.B. and Johnson,K.S. (1988). *Gene* **67**, 31–40.
30 Itoh,N., Mizumoto,K. and Kaziro,Y. (1984) *J. Biol. Chem.* **259**, 13923–13929.
31 Cong,P. and Shuman,S. (1993) *J. Biol. Chem.* **268**, 7256–7260.
32 Fresco,L.D. and Buratowski,S. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6624–6628.
33 Schwer,B. and Shuman,S. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4328–4332.
34 Niles,E.G., Condit,R.C., Caro,P., Davidson,K., Matusick,L. and Seto,J. (1986) *Virology* **153**, 96–112.
35 Upton,C., Stuart,D. and McFadden,G. (1991) *Virology* **183**, 773–777.
36 Pena,L., Yanez,R.J., Revilla,Y., Vinuela,E. and Salas,M.L. (1993) *Virology* **193**, 319–328.
37 Wang,P.W., Deng,L., Ho,C.K. and Shuman,S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 9573–9578.
38 Yue,Z., Maldonado,E., Pillutla,R., Cho,H., Reinberg,D. and Shatkin,A. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 12898–12903.
39 Denu,J.M., Stuckey,J.A., Saper,M.A. and Dixon,J.E. (1996) *Cell* **87**, 361–364.
40 McCracken,S., Fong,N., Rosonina,E., Yankulov,K., Brothers,G., Siderovski,D., Hessel,A., Foster,S., Amgen EST Program, Shuman,S. and Bentley,D.L. (1977) *Genes Dev.* **11**, 3306–3318.