THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 16, Issue of April 17, pp. 9577–9585, 1998

Printed in U.S.A.

The Guanylyltransferase Domain of Mammalian mRNA Capping Enzyme Binds to the Phosphorylated Carboxyl-terminal Domain of RNA Polymerase II*

(Received for publication, December 10, 1997, and in revised form, January 30, 1998)

C. Kiong Ho‡, Verl Sriskanda‡, Susan McCracken§, David Bentley§, Beate Schwer¶, and Stewart Shuman‡‖

From the ‡Molecular Biology Program, Sloan-Kettering Institute, New York, New York 10021, §Amgen Institute and Ontario Cancer Institute, Toronto, Ontario M5G 2C1, Canada, and the ¶Microbiology Department, Cornell University Medical College, New York, New York 10021

We have conducted a biochemical and genetic analysis of mouse mRNA capping enzyme (Mce1), a bifunctional 597-amino acid protein with RNA triphosphatase and RNA guanylyltransferase activities. The principal conclusions are as follows: (i) the mammalian capping enzyme consists of autonomous and nonoverlapping functional domains; (ii) the guanylyltransferase domain Mce1(211–597) is catalytically active *in vitro* and functional *in vivo* in yeast in lieu of the endogenous guanylyltransferase Ceg1; (iii) the guanylyltransferase domain *per se* binds to the phosphorylated RNA polymerase II carboxyl-terminal domain (CTD), whereas the triphosphatase domain, Mce1(1–210), does not bind to the CTD; and (iv) a mutation of the active site cysteine of the mouse triphosphatase elicits a strong growth-suppressive phenotype in yeast, conceivably by sequestering pre-mRNA ends in a nonproductive complex or by blocking access of the endogenous yeast triphosphatase to RNA polymerase II. These findings contribute to an emerging model of mRNA biogenesis wherein RNA processing enzymes are targeted to nascent polymerase II transcripts through contacts with the CTD. The phosphorylation-dependent interaction between guanylyltransferase and the CTD is conserved from yeast to mammals.

mRNA capping occurs by a series of three enzymatic reactions in which the 5′ triphosphate terminus of the primary transcript is cleaved to a diphosphate by RNA triphosphatase, capped with GMP by RNA guanylyltransferase, and methylated at the N-7 position of guanine by RNA (guanine 7) methyltransferase (1). *In vivo*, the capping reactions occur cotranscriptionally, *i.e.* the substrates for the capping enzymes are nascent RNA chains engaged within RNA polymerase II (pol II)[1] elongation complexes. There must exist a mechanism to target cap formation *in vivo* to transcripts made by pol II, because the capping enzymes have no inherent specificity for modifying pre-mRNAs *in vitro*. We and others (2–4) have suggested that targeting is achieved through direct physical interaction of one or more components of the capping apparatus with the phosphorylated carboxyl-terminal domain (CTD) of the largest subunit of pol II. This model is supported by the finding that recombinant yeast guanylyltransferase and methyltransferase proteins bind specifically and independently to the phosphorylated CTD *in vitro* (2). Are such direct protein-protein interactions conserved in higher eukaryotes? Does the triphosphatase component of the capping apparatus also interact with the CTD?

We know that the physical and functional organizations of the triphosphatase and guanylyltransferase components of the capping apparatus have diverged in fungi *versus* metazoans. The guanylyltransferases of *Saccharomyces cerevisiae* (Ceg1; 459 amino acids), *Schizosaccharomyces pombe* (Pce1; 402 amino acids), and *Candida albicans* (Cgt1; 449 amino acids) are monofunctional polypeptides that cap diphosphate-terminated RNAs (5–7). Transfer of GMP from GTP to the 5′ diphosphate terminus of RNA occurs in a two-stage reaction involving a covalent enzyme-GMP intermediate (8). The GMP is linked to the enzyme through a phosphoamide (P–N) bond to the ϵ-amino group of a lysine residue within a conserved motif, KXDG, found in all known cellular and DNA virus-encoded capping enzymes (9). The fungal guanylyltransferases display ~38% amino acid sequence identity overall. They are also functionally homologous, insofar as *PCE1* and *CGT1* can complement lethal *ceg1* mutations in *S. cerevisiae* (6, 7). The *S. cerevisiae* RNA triphosphatase is a 549-amino acid polypeptide encoded by the *CET1* gene (10). The Ceg1 and Cet1 polypeptides interact *in vivo* and *in vitro*.

Metazoan capping enzymes are bifunctional polypeptides with triphosphatase and guanylyltransferase activities (11, 12). Yagi *et al.* (12) isolated triphosphatase and guanylyltransferase domain fragments of the *Artemia salina* capping enzyme by partial proteolysis with trypsin. However, it was not clear from this work whether the functional domains overlapped structurally. The first metazoan-capping enzyme gene was isolated recently from *Caenorhabditis elegans* (13, 14). The 573-amino acid nematode protein consists of a carboxyl-terminal domain homologous to yeast Ceg1 and an amino-terminal domain that has strong similarity to the superfamily of protein phosphatases that act via a covalent phosphocysteine intermediate.

We recently isolated a mouse cDNA encoding the mammalian homologue of the *C. elegans* capping enzyme (2). The 597-amino acid mouse capping enzyme (Mce1) also consists of an amino-terminal phosphatase domain and a carboxyl-terminal guanylyltransferase domain. Here, we report that Mce1 is functional *in vitro* and *in vivo*. An autonomous RNA triphosphatase domain (amino acids 1–210) and an autonomous guanylyltransferase domain (amino acids 211–597) have been purified and characterized. We found that the guanylyltransferase domain *per se* binds to the pol II CTD

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed.

[1] The abbreviations used are: pol, polymerase; CTD, carboxyl-terminal domain; PAGE, polyacrylamide gel electrophoresis; DTT, dithiothreitol; GST, glutathione *S*-transferase; 5-FOA, 5-fluoroorotic acid.

and that this interaction requires CTD phosphorylation. The triphosphatase domain of mouse capping enzyme does not bind to the CTD. The phosphorylation-dependent interaction between guanylyltransferase and the CTD is conserved from yeast to mammals and provides a general mechanism for targeting caps to nascent pol II transcripts *in vivo*.

MATERIALS AND METHODS

Yeast Expression Plasmids—The *MCE1* cDNA was cloned into a customized yeast expression vector, pYX1-His, a derivative of pYX132 (*CEN TRP1*) in which six consecutive histidine codons and a unique *Nde*I site are inserted between the *Nco*I and *Bam*HI sites of pYX132 (pYX-132 was purchased from Novagen). The DNA insert of the resulting plasmid pYX1-MCE1 extended from the Mce1 translation start codon to an *Xho*I site located in the 3′ UTR. pYX1-MCE1 encodes the full-length 597-amino acid Mce1 polypeptide fused in-frame with an amino-terminal 12-amino acid leader peptide (MGSHHHHHHSGH). *MCE1* expression in this plasmid was under the control of a constitutive yeast *TPI* promoter. Amino-terminal deletion mutants of *MCE1* were constructed by PCR amplification with mutagenic primers that introduced an *Nde*I restriction site and a methionine codon in lieu of the codons for Ser-210 or Gln-259 or an *Nde*I site at Met-276. The PCR products were digested with *Nde*I and *Bgl*II and then inserted into pYX1-MCE1 so as to replace a 1.0-kilobase pair *Nde*I-*Bgl*II fragment of the wild type *MCE1* gene with restriction fragments encoding deleted Mce1 polypeptides. The mutated genes were named according to the amino acid coordinates of their polypeptide products, *i.e. MCE1(211–597)*, *MCE1(260–597)*, and *MCE1(276–597)*. Alanine-substitution mutations (C126A and K294A) in *MCE1* were programmed by synthetic oligonucleotides using the two-stage overlap extension method (15). pYX1-*MCE1* served as the template for the first round of amplification. The second-stage PCR products of the C126A and K294A reactions were digested with *Nde*I and *Bgl*II or *Nde*I and *Hin*dIII, respectively, and then ligated into the corresponding sites in pYX1-MCE1 in place of the wild type fragments. The presence of the desired mutations was confirmed in every case by dideoxy sequencing. We sequenced the entire restriction fragment insert in each pXY1 plasmid to exclude the introduction of unwanted mutations. A *CEN TRP1* vector, pYX1-CEG1, encoding yeast Ceg1 under the control of a *TPI* promoter, was constructed by inserting an *Nco*I-*Hin*dIII fragment of plasmid pGYCE-360 (16) into pYX132.

Test of MCE1 Function by Plasmid Shuffle—Strain YBS2 (*MAT***a** *ura3 trp1 lys2 leu2 ceg1::hisG* pGYCE-360), which is deleted at the chromosomal *CEG1* locus, is viable when it maintains an extrachromosomal copy of *CEG1* on a *CEN URA3* plasmid (pGYCE-360) (16). YBS2 was transformed with pXY1-based plasmids bearing wild type and mutant alleles of *MCE1*. Trp+ transformants were selected on medium lacking tryptophan. Individual colonies were patched on medium lacking tryptophan. Cells from single patches were then streaked on medium containing 0.75 mg/ml of 5-fluoroorotic acid (5-FOA). The plates were incubated at 30 °C. Mutations scored as lethal were those that did not support colony formation after 7 days. Individual colonies of the viable *MCE1* alleles were picked from the 5-FOA plate and patched to plates lacking tryptophan. All 5-FOA survivors were confirmed to be Ura−.

*Bacterial Expression Plasmids—Nde*I-*Xho*I fragments of pYX1-MCE1, pYX1-MCE1(211–597), pYX1-MCE1(260–597), and pYX1-MCE1(276–597) were inserted into to the T7-based expression plasmid pET16b (Novagen). The carboxyl-terminal deletion mutant *MCE1(1–210)* was constructed by PCR amplification from an MCE1 template with an antisense primer that introduced a translation stop at codon 211 and a flanking *Xho*I restriction site. The PCR product was digested with *Nde*I and *Xho*I and inserted into pET16b.

Expression and Purification of Recombinant Mce1 Protein—A 25-ml culture of *E. coli* BL21(DE3)/pET-MCE1 was grown at 37 °C in Luria-Bertani medium containing 0.1 mg/ml ampicillin until the A_{600} reached 0.5. The culture was adjusted to 0.4 mM isopropyl-β-D-thiogalactopyranoside, and incubation was continued at 17 °C for 17 h. Cells were harvested by centrifugation, and the pellet was stored at −80 °C. All subsequent procedures were performed at 4 °C. Thawed bacteria were resuspended in 5 ml of Buffer A (50 mM Tris-HCl, pH 7.5, 0.2 M NaCl, 10% sucrose). Lysozyme was added to a final concentration of 50 μg/ml; the suspension was incubated on ice for 10 min and then sonicated for 30 s. Triton X-100 was added to a final concentration of 0.1%, and sonication was repeated to reduce the viscosity of the lysate. Insoluble material was removed by centrifugation for 45 min at 18,000 rpm in a Sorvall SS34 rotor. The soluble extract was applied to a 0.5-ml column of Ni-NTA-agarose (Qiagen) that had been equilibrated with Buffer A containing 0.1% Triton X-100. The column was washed with the same buffer and then eluted step-wise with Buffer B (50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 10% glycerol) containing 50, 100, 200, 500, and 1000 mM imidazole. The polypeptide compositions of the column fractions were monitored by SDS-polyacrylamide gel electrophoresis (PAGE). The recombinant Mce1 polypeptide was retained on the column and recovered in the 200 and 500 mM imidazole eluates. These two fractions were pooled and dialyzed against Buffer C (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 2 mM DTT, 10% glycerol, 0.05% Triton X-100).

Expression and Purification of Mce1(1–210) and Mce1(211–597)—500-ml cultures of *E. coli* BL21(DE3) harboring pET-MCE1(1–210) or pET-MCE1(211–597) were grown at 37 °C until the A_{600} reached 0.5. The cultures were adjusted to 0.4 mM isopropyl-β-D-thiogalactopyranoside and incubated for 3 h at 37 °C. Cells were harvested by centrifugation and stored at −80 °C. Thawed bacteria were resuspended in 25 ml of Buffer A. Soluble lysates were prepared as described in the preceding section and then applied to 4-ml columns of Ni-NTA-agarose, which were eluted step-wise with 50, 100, 200, 500, and 1000 mM imidazole. The Mce1(1–210) and Mce1(211–597) polypeptides were retained on the Ni-NTA-agarose columns and recovered in the 200 and 500 mM imidazole eluates. The preparations were dialyzed against Buffer C. Enzyme fractions were stored at −80 °C and thawed on ice just prior to use. Protein concentrations were determined using the Bio-Rad dye binding assay with bovine serum albumin as a standard.

Glycerol Gradient Sedimentation—Aliquots (200 μl) of the dialyzed Ni-agarose preparations of Mce1, Mce1(1–210), and Mce1(211–597) were applied to a 4.8-ml 15–30% glycerol gradients containing 50 mM Tris-HCl (pH 8.0), 0.3 M NaCl, 2 mM DTT, 1 mM EDTA, and 0.1% Triton X-100. The gradients were centrifuged at 50,000 rpm for 24 h at 4 °C in a Beckman SW50 rotor. Fractions (0.2 ml) were collected from the bottom of the tube. Marker proteins catalase, bovine serum albumin, soybean trypsin inhibitor, and cytochrome C were sedimented in a parallel gradient.

Enzyme-GMP Complex Formation—Standard reaction mixtures (20 μl) containing 50 mM Tris-HCl (pH 8.0), 5 mM DTT, 5 mM $MgCl_2$, [α^{32}P]GTP as specified, and enzyme were incubated for 5 min at 37 °C. The reactions were halted by adding SDS to a final concentration of 1%. The samples were electrophoresed through a 12% polyacrylamide gel containing 0.1% SDS. Enzyme-[^{32}P]GMP complexes were visualized by autoradiographic exposure of the dried gel and was quantitated by scanning the gel with a FUJIX BAS1000 Bio-Imaging Analyzer.

RNA Triphosphatase Assay—RNA triphosphatase activity was assayed by liberation of $^{32}P_i$ from γ-^{32}P-labeled triphosphate-terminated poly(A) (17). Standard reaction mixtures (10 μl) containing 50 mM Tris-HCl (pH 7.5), 5 mM DTT, 10 pmol (of triphosphate termini) of [γ-^{32}P]poly(A), and enzyme as specified were incubated for 15 min at 37 °C. Reactions was halted by addition of 1 μl of 12 N formic acid. Aliquots of the mixtures were applied to a polyethyleneimine-cellulose TLC plate that was developed with 0.75 M potassium phosphate (pH 4.3). The release of $^{32}P_i$ from [γ-^{32}P]poly(A) was quantitated by scanning the TLC plate with a FUJIX BAS1000 Bio-Imaging Analyzer.

CTD Affinity Chromatography—Recombinant glutathione *S*-transferase (GST)-CTD, consisting of glutathione-S-transferase fused to the first 15 tandem heptad repeats from the mouse pol II CTD (consensus sequence YSPTSPS), was purified and coupled to glutathione agarose as described (2). GST-CTD was phosphorylated *in vitro* by incubation with HeLa cell extract and ATP (2). The phosphorylation reaction resulted in the addition of approximately 3 phosphates per molecule of GST-CTD. The GST-CTD-PO4 polypeptide was adsorbed to glutathione agarose and the resin was washed with buffer containing 1 M NaCl prior to use in affinity chromatography. The GST-CTD and GST-CTD-PO4 resins contained 3.0 mg of fusion protein per ml of agarose. Affinity chromatography was performed by mixing 25 μl of GST-CTD or GST-CTD-PO4 resins with 180–200-μl samples of Mce1 proteins in binding buffer containing 20 mM HEPES (pH 7.9), 0.1 mM EDTA, 1 mM DTT, 20% glycerol, 0.1 M NaCl, 0.5 μM microcystine, 1 mM β-glycerophosphate, 0.1% Nonidet P-40. After mixing for 1 h at 4 °C, the agarose beads were collected by centrifugation, and the supernatant was removed. The resins were washed three or four times with 400–500 μl of binding buffer, and then the bound material was eluted with 100 μl of binding buffer containing 0.9 M NaCl. Aliquots of the input sample and the 0.9 M NaCl eluates were analyzed by SDS-PAGE.

RESULTS

Mammalian Capping Enzyme—We have isolated a mouse cDNA that encodes Mce1, a putative mRNA capping enzyme

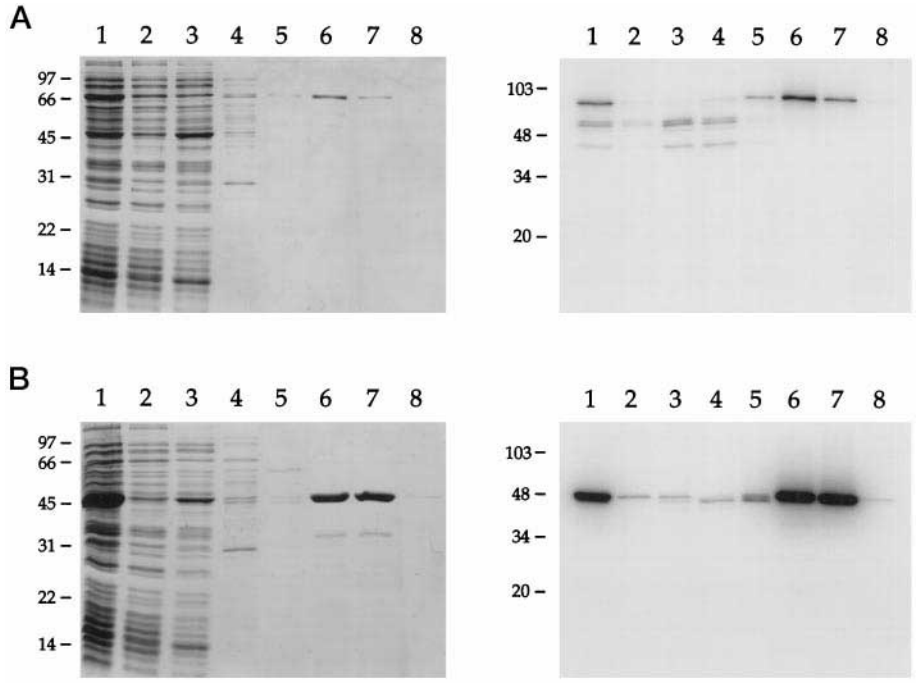

FIG. 1. **Purification and guanylyltransferase activity of recombinant Mce1 and Mce1(211–597).** The polypeptide compositions of recombinant Mce1 (*A, left panel*) and Mce1(211–597) (*B, left panel*) during Ni-agarose purification were analyzed by SDS-PAGE: *lane 1*, soluble bacterial lysate; *lane 2*, Ni-agarose flow-through; *lane 3*, Ni-agarose wash; *lane 4*, 50 mM imidazole eluate; *lane 5*, 100 mM imidazole eluate; *lane 6*, 200 mM imidazole eluate; *lane 7*, 500 mM imidazole eluate; *lane 8*, 1 M imidazole eluate. The gels were fixed and stained with Coomassie Blue dye. The positions and sizes (in kDa) of coelectrophoresed marker polypeptides are indicated. The guanylyltransferase activity of recombinant Mce1 (*A, right panel*) and Mce1(211–597) (*B, right panel*) during Ni-agarose purification was assayed by enzyme-GMP complex formation. Reaction mixtures containing 50 mM Tris-HCl (pH 8.0), 5 mM DTT, 5 mM $MgCl_2$, 0.17 μM [α-^{32}P]GTP, and 1 μl of the protein fractions specified above in *lanes 1–8* were incubated for 5 min at 37 °C. The reaction products were analyzed by SDS-PAGE. Autoradiographs of the gels are shown. The positions and sizes (in kDa) of prestained marker polypeptides are indicated.

(GenBank™ accession no. AF034568). The size of the Mce1 polypeptide (597-amino acids; 68 kDa) agrees with the size of the 68-kDa enzyme-GMP complex formed by the capping enzyme isolated from mouse cell extracts (18). The amino acid sequence of Mce1 suggests that it is a bifunctional enzyme consisting of an amino-terminal phosphatase domain and a C-terminal guanylyltransferase domain. In this respect, it resembles the *C. elegans* capping enzyme (13, 14), to which it displays 43% sequence identity. The carboxyl-terminal portion of Mce1 contains the defining sequence motifs of the covalent nucleotidyl transferase superfamily (9), including the **K***X*-DG*X*R sequence that constitutes the active site. The lysine side chain within this motif (Lys-294 in Mce1) reacts with GTP to form a covalent enzyme-GMP intermediate. The amino-terminal portion of Mce1 contains the (I/V)H**C***X*AG*X*GR(S/T)G signature motif of the dual-specificity protein phosphatase/protein tyrosine phosphatase enzyme family. These proteins catalyze phosphoryl transfer from a protein phosphomonoester substrate to the thiol of a cysteine on the enzyme to form a covalent phosphocysteine intermediate, which is then attacked by water to liberate phosphate (19). The cysteine within the signature motif is the active site of phosphoryl transfer and is thus essential for reaction chemistry. Cys-126 is predicted to be the active site of phosphoryl transfer by Mce1.

Catalytic Activity of Mouse Capping Enzyme—The biochemical properties of the mammalian capping enzyme were examined after expressing Mce1 and its component domains in bacteria. First, the full-length *MCE1* coding sequence was inserted into an inducible T7 RNA polymerase-based pET vector such that a histidine-rich amino-terminal leader (His-tag) was fused to the 597-amino acid Mce1 polypeptide. The pET-Mce1 plasmid was introduced into *E. coli* BL21(DE3), a strain that contains the T7 RNA polymerase gene under the control of a *lac*UV5 promoter. A 68-kDa polypeptide corresponding to His-tagged Mce1 was detected by SDS-PAGE in whole-cell and soluble lysates of isopropyl-β-D-thiogalactopyranoside-induced bacteria (Fig. 1*A* and data not shown). To assay guanylyltransferase activity of the expressed Mce1 protein, we incubated soluble protein from induced bacteria in the presence of [α-^{32}P]GTP and a divalent cation. This resulted in the formation of a ^{32}P-labeled nucleotidyl-protein adduct that migrated as a 68-kDa species during SDS-PAGE (Fig. 1*A*). Three other polypeptides in the range of 45–55 kDa were labeled with GMP to a lesser extent. We detected no label transfer from GTP to polypeptides of these sizes in extracts prepared from bacteria that lacked the *MCE1* gene (not shown). Purification of the His-tagged 68-kDa Mce1 guanylyltransferase was achieved by adsorption to Ni-agarose and elution with 200–500 mM imidazole (Fig. 1*A*). The imidazole eluate fractions were highly enriched with respect to Mce1, as judged by SDS-PAGE (Fig. 1*A, left panel*). Note that the lower molecular weight polypeptides with guanylyltransferase activity were recovered in the Ni-agarose flow-through and wash fractions (Fig. 1*A, right panel*), which suggested that these were proteolytic fragments of Mce1 that lacked the amino-terminal His-tag.

The Ni-agarose Mce1 preparation was centrifuged through a 15–30% glycerol gradient. The fractions were assayed for enzyme-GMP complex formation and for RNA triphosphatase. The latter activity was detected by the release of $^{32}P_i$ from γ^{32}P-labeled poly(A). A single peak of guanylyltransferase activity was coincident with the RNA triphosphatase activity profile (Fig. 2*A*). Mce1 sedimented at 3.9 *S* relative to marker proteins centrifuged in a parallel gradient. We surmise from this experiment that recombinant Mce1 is a bifunctional monomeric enzyme.

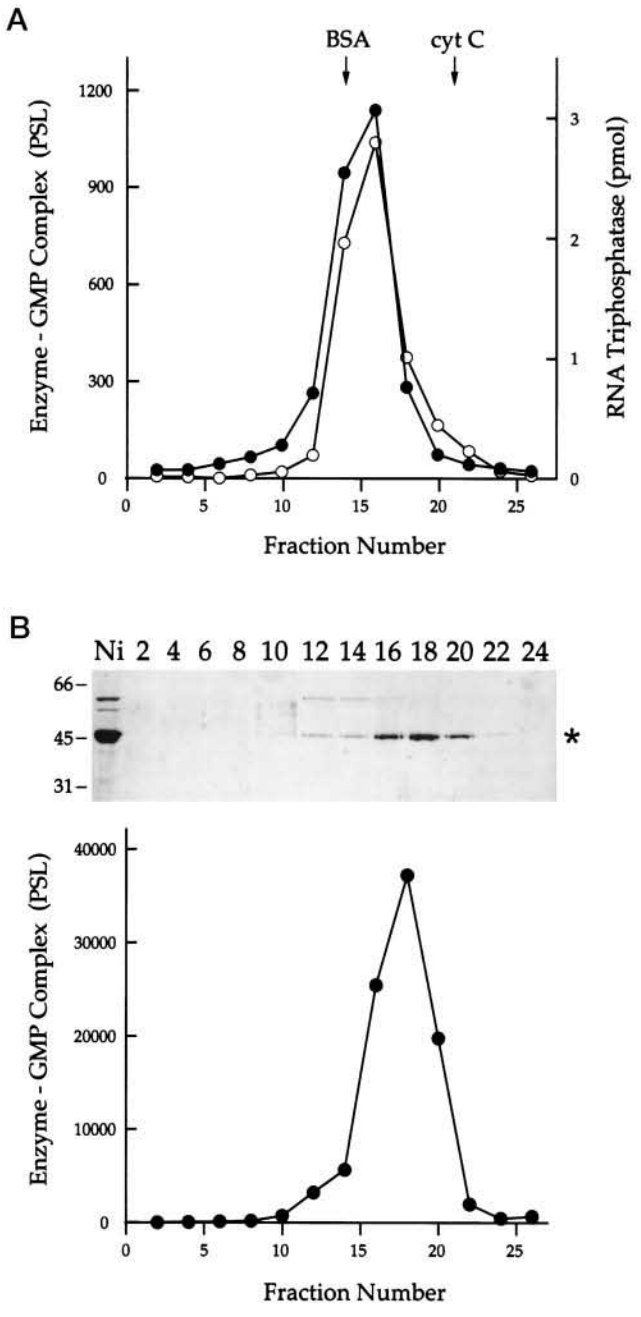

FIG. 2. **Glycerol gradient sedimentation of Mce1 and Mce1(211–597).** *A*, the Mce1 Ni-agarose preparation was sedimented in a 15–30% glycerol gradient and alternate fractions were assayed for enzyme-GMP complex formation (●) and RNA triphosphatase activity (○). The guanylyltransferase reaction mixtures contained 0.17 μM [α-^{32}P]GTP and 1 μl of the indicated fractions. Guanylyltransferase activity was gauged by the signal intensity of the radiolabeled 68-kDa Mce1 polypeptide (*PSL*, photo stimulatable luminescence), plotted on the *y* axis at left. The RNA triphosphatase reaction mixtures contained 10 pmol (of triphosphate termini) of [γ-^{32}P] poly(A) and 0.1 μl (1 μl of a 1:10 dilution) of the indicated gradient fractions. Incubation was for 15 min at 37 °C. Activity (in pmol of P_i released) is plotted on the *y* axis at right. The peaks of marker proteins bovine serum albumin and cytochrome *c*, centrifuged in a parallel gradient, are indicated by *arrows*. *B*, the Mce1(211–597) Ni-agarose preparation (*Ni*) was sedimented in a 15–30% glycerol gradient, and gradient fractions were analyzed by SDS-PAGE. Polypeptides were visualized by staining with Coomassie Blue dye. The position of Mce1(211–597) is denoted by an *asterisk* on the *right*. The position and sizes (in kDa) of marker polypeptides are indicated on the left. Guanylyltransferase reaction mixtures contained 0.17 μM [α-^{32}P]GTP and 1 μl of the indicated gradient fractions; incubation was for 5 min at 37 °C. Guanylyltransferase activity was gauged by the signal intensity of the radiolabeled 46-kDa Mce1(211–597) polypeptide (*PSL*).

Autonomous Guanylyltransferase and Triphosphatase Domains—To evaluate whether the carboxyl-terminal portion of Mce1 constituted an autonomous functional guanylyltransferase domain, we expressed the protein segment from residues 211–597 as a His-tagged fusion protein. The choice of residue 211 as a domain breakpoint was based on its location 84 amino acids from the putative Mce1 active site Lys-294; the active site lysine of the monofunctional *Chlorella* virus guanylyltransferase is located 82 amino acids from its amino terminus (20, 21). We found that isopropyl-β-D-thiogalactopyranoside-induced bacteria accumulated substantial amounts of a soluble 46-kDa polypeptide corresponding to Mce1(211–597), which reacted with [α-^{32}P]GTP to form a 47-kDa radiolabeled enzyme-GMP adduct (Fig. 1*B*). Mce1(211–597) adsorbed to Ni-agarose and was eluted at 200–500 mM imidazole (Fig. 1*B*). Mce1(211–597) sedimented as a discrete 2.8 *S* peak during glycerol gradient centrifugation (Fig. 2*B*). We surmise that Mce1 the guanylyltransferase component is a monomer in solution.

Induced expression of the amino-terminal portion of Mce1 from residues 1–210 as a His-tagged fusion protein resulted in the accumulation of a soluble 28-kDa polypeptide that could be adsorbed to Ni-agarose and eluted with 200–500 mM imidazole (Fig. 3*A*). Mce1(1–210) was recovered in nearly homogeneous form at this stage. The Ni-agarose preparation was highly active as an RNA triphosphatase (see below). As one might expect, Mce1(1–210) was incapable of forming a covalent adduct with [α-^{32}P]GTP (not shown). Mce1(1–210) sedimented in a glycerol gradient as a discrete peak of 2.5 *S* (Fig. 3*B*). The RNA triphosphatase activity profile coincided with the abundance of the 28-kDa Mce1(1–210) polypeptide. We conclude from these experiments that the mammalian capping enzyme is composed of autonomous nonoverlapping triphosphatase and guanylyltransferase domains.

Characterization of the RNA Triphosphatase Domain—The standard RNA triphosphatase reaction contained 50 mM Tris-HCl (pH 7.5) and 1 μM γ^{32}P-labeled poly(A). The extent of γ-phosphate hydrolysis during a 15-min incubation at 37 °C was proportional to input Mce1(1–210) (Fig. 4*A*). In the linear range of enzyme dependence, 1 pmol of $^{32}P_i$ was released per fmol of protein. The specific activity of the Ni-agarose preparation was 37 units/μg. (One unit of enzyme releases 1 nmol of $^{32}P_i$ from γ^{32}P poly(A) in 15 min.) A kinetic analysis showed that the initial rate of P_i release was proportional to enzyme concentration (Fig. 4*B*). Mce1(1–210) hydrolyzed 1.2–2 molecules of P_i/enzyme/s at steady state. RNA triphosphatase activity was optimal in 50 mM Tris buffer at pH 7.0–7.5; the extents of γ-phosphate release at pH 6.0 and pH 9.5 were 70 and 35%, respectively, of the activity seen at pH 7.5 (not shown). Activity was optimal in the absence of a divalent cation and was unaffected by EDTA. Inclusion of divalent cations elicited a concentration-dependent inhibition of RNA triphosphatase activity (Fig. 4*C*). 75% inhibition was observed at 0.5 mM $MgCl_2$ or $MnCl_2$. Mce1(1–210) did not catalyze release of $^{32}P_i$ from [γ-^{32}P]ATP (not shown).

Characterization of the Guanylyltransferase Domain—Formation of the covalent enzyme-GMP complex by Mce1(211–597) was dependent on the concentration of GTP from 0.01 to 2 μM (Fig. 5*A*). The yield of guanylated enzyme plateaued at 5–10 μM GTP; we estimated that at least 30% of the enzyme molecules were labeled with GMP *in vitro*. (It has been our experience that a significant fraction of any recombinant guanylyltransferase purified from bacteria is already guanylylated and thus cannot be labeled during the *in vitro* reaction.) Enzyme-GMP complex formation was strictly dependent on a divalent cation cofactor. Either magnesium or manganese sufficed; ac-

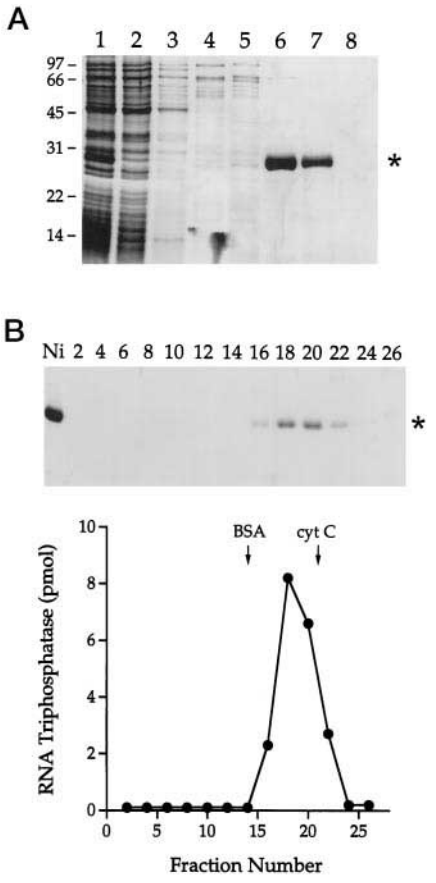

FIG. 3. **Purification and RNA triphosphatase activity of Mce1(1–210).** *A*, the polypeptide composition of Mce1(1–210) protein during Ni-agarose purification was analyzed by SDS-PAGE: *lane 1*, soluble bacterial lysate; *lane 2*, Ni-agarose flow-through; *lane 3*, Ni-agarose wash; *lane 4*, 50 mM imidazole eluate; *lane 5*, 100 mM imidazole eluate; *lane 6*, 200 mM imidazole eluate; *lane 7*, 500 mM imidazole eluate; *lane 8*, 1 M imidazole eluate. A Coomassie Blue-stained gel is shown. The Mce1(1–210) polypeptide is denoted by an *asterisk*. *B*, glycerol gradient sedimentation. The Mce1(1–210) Ni-fraction preparation was sedimented in a 15–30% glycerol gradient. The polypeptide compositions of the indicated fractions were analyzed by SDS-PAGE. A Coomassie Blue-stained gel is shown. Aliquots of the glycerol gradient fractions were assayed for RNA triphosphatase activity. The reaction mixtures contained 10 pmol (of triphosphate termini) of [γ-^{32}P]poly(A) and 1 μl of a 1:1000 dilution of the indicated gradient fractions. Incubation was for 15 min at 37 °C.

tivity was optimal at 1–2 mM $MnCl_2$ or 5–10 mM $MgCl_2$ (Fig. 5*B*). The rate of enzyme-guanylate formation was slightly greater in the presence of manganese *versus* magnesium (Fig. 5*C*). The Ni-agarose preparation of Mce1(211–597) displayed trace levels of RNA triphosphatase activity. The specific activity (0.02 units/μg of protein) was 0.05% that of Mce1(1–210). The residual triphosphatase was resolved from the guanylyltransferase activity during glycerol gradient sedimentation of Mce1(211–597), suggesting that this phosphatase was an *E. coli* contaminant. We surmise that the guanylyltransferase domain of Mce1 has no triphosphatase activity *per se*.

To confirm that the Mce1 domains actually catalyzed cap formation on triphosphate RNA ends, we incubated Mce1(1–210) and Mce1(211–597) together with unlabeled triphosphate-terminated poly(A) in the presence of [α-^{32}P]GTP and 0.1 mM $MnCl_2$. The RNA was recovered free of GTP by several rounds of TCA precipitation, and then deproteinized and ethanol precipitated. Digestion of the labeled RNA with nuclease P1 liberated a single labeled species that migrated during PEI cellulose thin layer chromatography with GpppA (not shown).

Mce1 Activity in Vivo—To test whether the mouse capping enzyme was functional *in vivo*, we examined the ability of the wild type full-length *MCE1* cDNA to complement a *ceg1* null mutation by using the plasmid shuffle technique (6, 16). *MCE1* was cloned into a yeast *CEN TRP1* expression plasmid under the control of a constitutive yeast *TPI* promoter. The *MCE1* plasmid was introduced into YBS2, a yeast strain in which the chromosomal *CEG1* locus is deleted and that is dependent for growth on maintenance of an extrachromosomal copy of *CEG1* on a *CEN URA3* plasmid. Trp+ transformants were plated on medium containing 5-FOA to select against retention of the wild type *CEG1* gene. Cells bearing the *TRP1 MCE1* plasmid grew readily on 5-FOA, whereas cells containing the *TRP1* vector plasmid were incapable of growth on 5-FOA. A mutated allele, *MCE1-K294A*, in which the active site nucleophile of the guanylyltransferase Lys-294 was substituted by alanine, was completely incapable of supporting yeast growth in the plasmid shuffle assay (Fig. 6*A*). We conclude that the mammalian capping enzyme is functional as a guanylyltransferase *in vivo*.

Cells transformed with *MCE1(211–597)* were also able to form colonies on 5-FOA, signifying that the guanylyltransferase domain by itself could replace Ceg1 *in vivo*. We noted that FOA-resistant *MCE1(211–597)* colonies arose with a slight delay compared with *MCE1* transformants. The FOA-selected *MCE1* and *MCE1(211–597)* isolates were tested for growth on rich medium in parallel with a strain bearing *CEG1* in the *TPI*-based vector. *MCE1* cells grew as well as *CEG1* cells; however, *MCE1(211–597)* cells grew somewhat more slowly at 30 °C, as gauged by colony size (Fig. 6*B*). Growth of *MCE1* and *MCE1(211–597)* cells was unaffected at 37 °C (not shown).

To define the amino-terminal boundary of the guanylyltransferase domain, we constructed truncated alleles *MCE1(260–597)* and *MCE1(276–597)*. Neither allele was capable of sustaining growth of the *ceg1* strain on 5-FOA (Fig. 6*A*). Expression of Mce1(260–597) in bacteria resulted in the production of predominantly insoluble protein; what little soluble recombinant protein was obtained was inactive in enzyme-GMP complex formation (not shown). We infer that amino-terminal truncation to 35 amino acids upstream of the active site Lys-294 abrogates mouse guanylyltransferase activity. This is in accord with the observations that an amino-terminal deletion of the yeast guanylyltransferase to 32 amino acids upstream of its active site was lethal *in vivo* and that the recombinant truncated yeast protein was catalytically inactive *in vitro* (16).

Interaction of Mce1 with the Pol II CTD—We used affinity chromatography to analyze the binding of mouse capping enzyme and its constituent domains to the pol II CTD. The test samples were incubated in parallel with a glutathione-Sepharose resin containing an immobilized ligand composed of either (i) glutathione *S*-transferase fused to the first 15 tandem heptad repeats of the mouse pol II CTD (consensus sequence YSPTSPS) or (ii) GST-CTD fusion protein that had been phosphorylated *in vitro* using HeLa nuclear extract as a source of kinase activity. The phosphorylation reaction resulted in the addition of approximately 3 phosphates per molecule of GST-CTD. After phosphorylation *in vitro*, the GST-CTD resin was washed extensively with 1 M NaCl to remove residual non-CTD proteins before performing affinity chromatography with mouse capping enzyme. The 68-kDa full-length Mce1 protein (residues 1–597) was prepared by *in vitro* translation in the presence of [^{35}S]methionine, and the translation product was subjected to affinity chromatography. We found that ^{35}S-labeled Mce1 bound specifically to the phosphorylated CTD (Fig. 7*A, lane P*) but not to the nonphosphorylated CTD control (*lane C*). The 46-kDa Mce1(211–597) guanylyltransferase domain by itself also bound specifically to phosphorylated CTD (Fig. 7*A*).

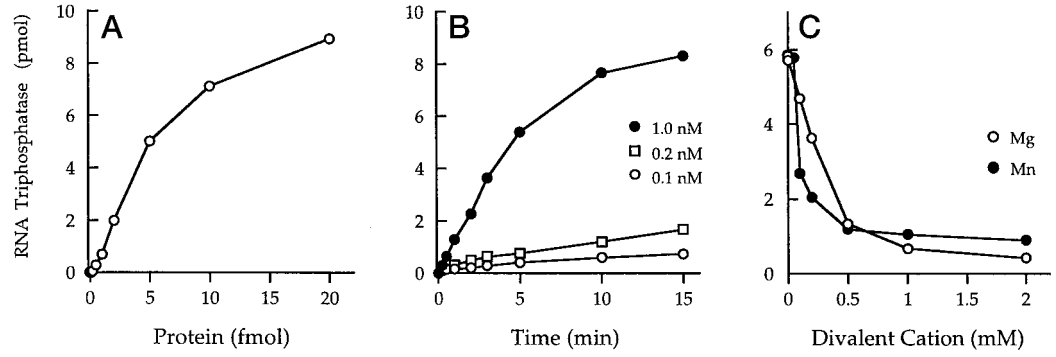

FIG. 4. **Characterization of the RNA triphosphatase domain.** *A*, protein titration. Reaction mixtures (10 μl) containing 50 mM Tris-HCl (pH 7.5), 5 mM DTT, 10 pmol (of triphosphate termini) of [γ-^{32}P]poly(A), and the indicated amounts of Mce1(1–210) (Ni-agarose fraction) were incubated for 15 min at 37 °C. P_i release is plotted as a function of input protein. *B*, kinetics. Reaction mixtures contained 1 μM [γ-^{32}P]poly(A) and 0.1, 0.2, or 1 nM Mce1(1–210). The reaction was initiated by the addition of enzyme. Aliquots (10 μl) were withdrawn at the times indicated and quenched with formic acid. P_i release is plotted as function of incubation time. *C*, inhibition by divalent cations. Reaction mixtures (10 μl) containing 10 pmol of [γ-^{32}P]poly(A), 5 fmol of Mce1(1–210), and $MgCl_2$ or $MnCl_2$ as indicated were incubated for 15 min at 37 °C. P_i release is plotted as a function of divalent cation concentration.

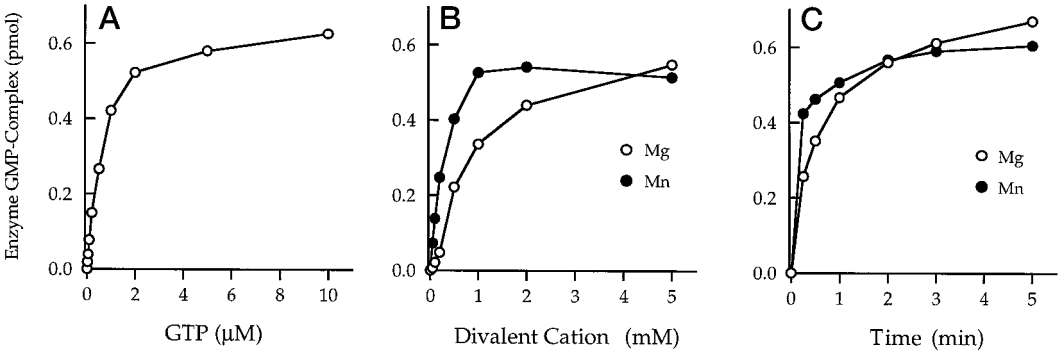

FIG. 5. **Analysis of enzyme-GMP formation by Mce1(211–597).** *A*, GTP dependence. Reaction mixtures (20 μl) containing 50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 5 mM DTT, 2 pmol of Mce1(211–597) (Ni-agarose fraction), and [α-^{32}P]GTP as indicated were incubated for 5 min at 37 °C. The products were analyzed by SDS-PAGE. The amount of ^{32}P-GMP (pmol) transferred to Mce1(211–507) is plotted as a function of GTP concentration. *B*, divalent cation dependence. Reaction mixtures (20 μl) containing 5 μM [α-^{32}P]GTP, 2 pmol of Mce1(211–597), and $MgCl_2$ or $MnCl_2$ as indicated were incubated for 5 min at 37 °C. Enzyme GMP-complex formation is plotted as a function of divalent cation concentration. *C*, kinetics. Reaction mixtures contained (per 20 μl) 5 μM [α-^{32}P]GTP, 2 pmol of Mce1(211–597), and either 5 mM $MgCl_2$ or 5 mM $MnCl_2$. The reaction was initiated by the addition of enzyme. Aliquots (20 μl) were withdrawn at the times indicated. Enzyme GMP-complex formation is plotted as function of incubation time.

In contrast, the RNA triphosphatase domain Mce1(1–210) did not interact with either phosphorylated or nonphosphorylated CTD (Fig. 7*A*). We conclude that the triphosphatase domain has no intrinsic affinity for the CTD and that its recruitment to the pol II transcription complex is mediated via its connection in *cis* to the guanylyltransferase. Guanylyltransferase deletion mutants Mce1(260–597) and Mce1(276–597) were unable to bind the phosphorylated CTD (Fig. 7*A*). Apparently, deletion of the segment of Mce1 from residues 211–259 disrupts CTD-PO4 recognition as well as guanylyltransferase activity.

CTD recognition by the isolated guanylyltransferase domain was confirmed by performing affinity chromatography with recombinant proteins produced in *E. coli*. The 46-kDa Mce1(211–597) polypeptide bound the CTD-PO4 resin and was recovered in the high salt eluate. It did not bind to the control CTD resin (Fig. 7*B, upper panel*). Assay of the fractions for enzyme-GMP complex formation confirmed that the guanylyltransferase bound specifically to the phosphorylated CTD (Fig. 7*B, lower panel*). The recombinant triphosphatase domain Mce1(1–210) was not capable of binding to either CTD or CTD-PO4 (Fig. 7B).

Cis Negative Effect of an Active Site Mutation of the RNA Triphosphatase on Mce1 Function Yeast—Yeast cells bearing *MCE1-C126A*, an allele with an alanine substitution at the active site cysteine of the triphosphatase domain, formed very tiny colonies on 5-FOA plates, even after 6 days of incubation. This contrasted starkly with the behavior of *MCE1* cells, which formed colonies of normal size after 2 days on 5-FOA. *MCE1-C126A* cells displayed a severe growth defect when plated on rich medium at 30 °C (Fig. 6*B*) and were unable to grow on YPD plates at 37 °C (not shown). The growth suppressive effects of an inactivating point mutation in the triphosphatase were far more severe than a complete deletion of the triphosphatase domain (Fig. 6*B*). We hypothesize that the C126A mutation interferes with the endogenous yeast triphosphatase, either by sequestering the nascent pre-RNA ends in a nonproductive complex or by blocking access of the yeast triphosphatase to RNA polymerase II.

DISCUSSION

Functional Domains of Mammalian Capping Enzyme—Fungi and *Chlorella* virus encode monofunctional guanylyltransferases, whereas mammals encode a bifunctional triphosphatase-guanylyltransferase enzyme. Here, we demonstrate that the triphosphatase and guanylyltransferase domains of mammalian capping enzyme are distinct and nonoverlapping. The guanylyltransferase domain is a 46-kDa monomer that extends from Mce1 residue 211 to residue 597. The recombinant domain is catalytically active *in vitro* and is capable of genetically complementing the function of the yeast guanylyltransferase Ceg1 *in vivo*. A recent study of the *C. elegans* enzyme Cel1 documented triphosphatase activity of an amino-

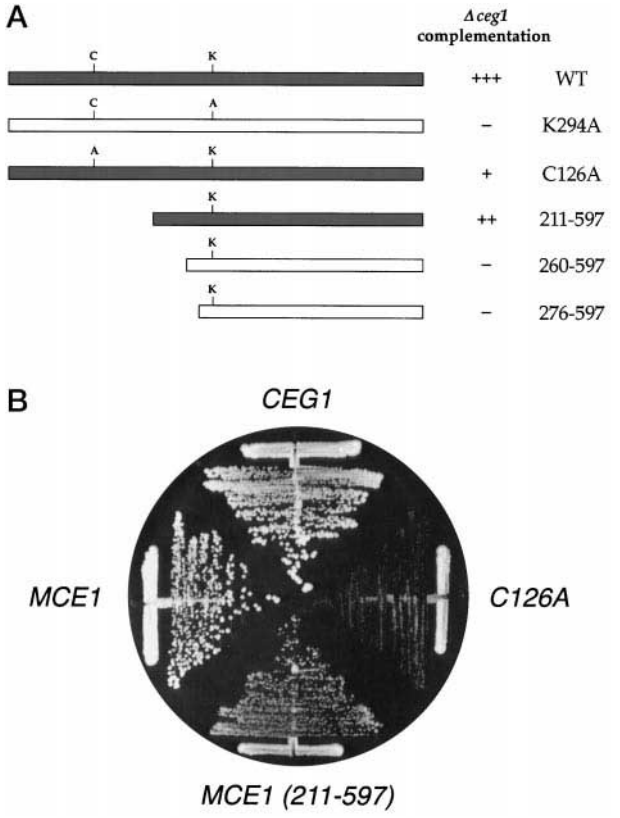

FIG. 6. **Complementation of Δ*ceg1* by *MCE1*.** *A*, the polypeptide encoded by wild type and mutant alleles of *MCE1* are depicted as horizontal bars with N termini at the *left* and C termini at the *right*. The positions of the active site cysteine (Cys-126; *C*) for RNA triphosphatase and the active site lysine (Lys-294; *K*) for the guanylyltransferase are indicated. Gene function was tested by plasmid shuffle. Complementation by plasmids containing *MCE1* alleles was gauged relative to a *CEG1* control plasmid. Alleles that supports the formation of wild type sized colonies after 3 days on 5-FOA are scored as +++. Alleles that yielded slightly smaller colonies are scored as ++, whereas those that required 7 days to form macroscopic colonies are denoted by +. Lethal mutations (–) were those that formed no colonies after 7 days. *WT*, wild type. *B*, YBS2 derivatives bearing the indicated capping enzyme genes under the control of a *TPI* promoter were streaked on YPD agar plates. The plates were photographed after incubation for 4 days at 30 °C.

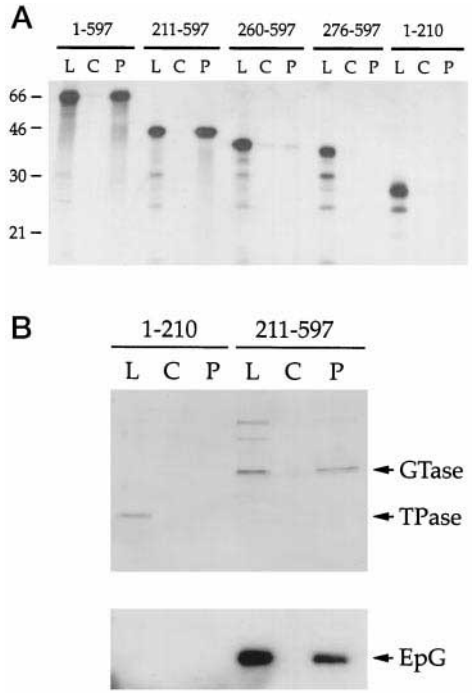

FIG. 7. **CTD affinity chromatography.** *A*, [^{35}S]methionine-labeled full-length Mce1 polypeptide (1–597) and truncated derivatives Mce1(211–597), Mce1(260–597), Mce1(276–597), and Mce1(1–210) were prepared by translation *in vitro* in a rabbit reticulocyte lysate. The transcripts encoding these proteins were transcribed *in vitro* by T7 RNA polymerase from pET-based vectors. CTD-affinity chromatography was performed as described under "Materials and Methods," and aliquots of the input fraction (*L*) (1.1% of total material loaded) and the high-salt eluates from the nonphosphorylated GST-CTD control resin (*C*) (10% of eluate) and the phosphorylated GST-CTD resin (*P*) (10% of eluate) were analyzed by SDS-PAGE. An autoradiograph of the gel is shown. The positions and sizes (in kDa) of prestained marker proteins are indicated on the *left*. *B*, partially purified recombinant Mce1(1–210) and Mce1(211–597) were subjected to CTD affinity chromatography. Aliquots of the input fraction (*L*) (5% of total material loaded) and the high-salt eluates from the nonphosphorylated GST-CTD control resin (*C*) (10% of eluate) and the GST-CTD-PO4 resin (*P*) (10% of eluate) were analyzed by electrophoresis through an 8–16% polyacrylamide gradient gel in 0.1% SDS. Polypeptides were visualized by silver staining. The positions of the guanylyltransferase (*GTase*) and triphosphatase (*TPase*) polypeptides are denoted by *arrows*. Aliquots (1 μl) of each fraction (representing 0.5% of the loaded material and 1% of the high salt eluates) were assayed for enzyme-GMP complex formation. The products were analyzed by SDS-PAGE; the labeled covalent intermediate (*EpG*) was detected by autoradiography of the gel (*lower panel*).

terminal domain (13), but the guanylyltransferase activity could not be demonstrated. We found that the proximal margin of the active guanylyltransferase domain is located 84 amino acids upstream of the presumptive active site nucleophile Lys-294. This is in keeping with the positions of the N termini of the yeast and *Chlorella* virus guanylyltransferases 67–82 residues upstream of their respective lysine nucleophiles (6, 7, 16, 20, 22). The fact that the *in vivo* activity of Mce1 is abrogated by the K294A mutation supports our designation of this residue as the active site. Substitutions of the equivalent residues of Ceg1 and Pce1 are also lethal *in vivo* (6, 16, 22). The function of the Mce1 guanylyltransferase domain is abolished when residues 211–259 are deleted. Deletion to a similar point of the yeast guanylyltransferase Ceg1 is also lethal (16). In the crystal structure of *Chlorella* virus guanylyltransferase, the analogous amino-terminal segment is located on the protein surface and consists of two antiparallel β-sheets followed by an α-helix (21). This region contains no residues that make direct contact with GTP in the enzyme-GTP cocrystal (21). The fact that recombinant Mce1(260–597) expressed in bacteria was predominantly insoluble complicated efforts to distinguish whether the loss of activity upon amino-terminal truncation was caused by defective protein folding or a primary defect in catalysis.

The RNA triphosphatase domain of Mce1 is contained within a segment consisting of residues 1–210. Using triphosphate-terminated poly(A) as a substrate, we determined a steady-state turnover number of 1–2 molecules of P_i release per second per enzyme. This value is quite close to the turnover number of 0.5–0.8 s^{-1} reported for the RNA triphosphatase component of the vaccinia capping enzyme (23), which acts via a different mechanism (see below). The Mce1 RNA triphosphatase activity is lower in the presence of divalent cations, which are essential for the guanylyltransferase reaction. However, because γ-phosphate cleavage is much faster than RNA guanylylation, an order of magnitude rate decrement in the triphosphatase step is unlikely to limit the overall rate of cap formation (1).

The finding that yeast *ceg1* cells expressing the full-length Mce1 protein grow slightly better than cells containing the mouse guanylyltransferase domain alone suggests that an active triphosphatase in *cis* enhances Mce1 function in yeast. A triphosphatase active site mutation C126A exerts a strong negative effect on yeast cell growth. The *CET1* gene encoding yeast RNA triphosphatase is essential (10). Assuming that the RNA triphosphatase activity of Cet1 is critical for cell growth,

then we interpret our results as indicating that (i) the mouse guanylyltransferase domain alone can cooperate with, or at least not interfere with, the function of the yeast triphosphatase, and (ii) the presence of a catalytically inert mouse triphosphatase linked to the catalytically active guanylyltransferase somehow blocks the action of the endogenous yeast enzyme.

Our findings concerning the domain structure of Mce1 agree with earlier reports that the guanylyltransferases isolated from rat liver and brine shrimp copurify with an RNA triphosphatase activity (11, 12). In the case of the brine shrimp protein, Yagi *et al.* (12) showed that both catalytic activities residue within a single 73-kDa polypeptide that was converted by partial proteolysis into catalytically active domains: a 20-kDa triphosphatase module could be separated from a 44-kDa fragment guanylyltransferase domain. The sizes of these two active fragments are consistent with those of the two autonomous domains of the mouse capping enzyme identified in the present study. It is therefore likely that bifunctional triphosphatase-guanylyltransferase enzymes are present in all higher eukaryotes.

Evolution of the Capping Apparatus—The linear order of amino-terminal RNA triphosphatase and carboxyl-terminal guanylyltransferase domains in Mce1 is superficially similar to the arrangement of the triphosphatase and guanylyltransferase active sites in the vaccinia capping enzyme (24–27). However, the amino acid sequences of the Mce1 and vaccinia amino-terminal segments are unrelated, and the properties of the vaccinia triphosphatase differ from those of the metazoan triphosphatases in two key respects: (i) lack of domain autonomy, and (ii) divalent cation dependence. The vaccinia triphosphatase and guanylyltransferase active sites are distinct, but the two activities are not partitioned into discrete and separable structural modules as they are in Mce1. Instead, moieties on the vaccinia enzyme that are essential for triphosphatase and guanylyltransferase activity overlap within a 545-amino acid polypeptide (23, 26, 27). The triphosphatase activity of the vaccinia capping enzyme depends absolutely on a divalent cation cofactor, whereas the rat liver and brine shrimp triphosphatases, as well as the recombinant *C. elegans* and mouse triphosphatases, require no divalent cation for activity (in fact, they are inhibited by divalent cations). The sequences of the *C. elegans* and mouse enzymes indicate that γ-phosphate cleavage occurs through a phosphoenzyme intermediate. Mutation of the active site cysteine of the *C. elegans* triphosphatase domain abolished enzyme activity (13). We have been unable to detect a phosphoenzyme intermediate for the vaccinia triphosphatase, which suggests that covalent catalysis does not apply. Mutational analysis of the vaccinia triphosphatase provides additional evidence for a distinct mechanism. We have identified four acidic side chains that are essential for catalysis by vaccinia triphosphatase and are conserved among the poxvirus and African swine fever virus capping enzymes (26, 27). One or more of these acidic residues is likely to bind the essential metal ion(s). The RNA triphosphatase isolated from *S. cerevisiae*, also depends completely on a divalent cation cofactor (28). The amino acid sequence of yeast Cet1 bears no similarity to the mammalian triphosphatase domains, but it does display local similarities to the vaccinia triphosphatase. We surmise that higher eukaryotes have diverged from vaccinia and yeast with respect to both mechanism and structure of the triphosphatase component of the capping machinery. In contrast, the reaction mechanism and structure of the guanylyltransferase components are conserved in yeasts, metazoans, and DNA viruses (14, 21).

Interaction of the Capping Apparatus with the CTD of RNA Polymerase II—Cap formation *in vivo* is targeted to the nascent chains synthesized by pol II. Placing a mammalian pol II transcription unit under the control of a pol III promoter results in a failure to cap the transcript (29). We have elaborated a solution to the problem of how pol II transcripts are specifically singled out for capping whereby the capping enzymes are targeted to pre-mRNA by binding to the phosphorylated CTD of the largest subunit of RNA polymerase II (2). The CTD, which is unique to pol II, consists of a tandem array of a heptapeptide repeat with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. The mammalian pol II large subunit has 52 tandem repeats, whereas the *S. cerevisiae* subunit has 27 copies. The pol II largest subunit exists in two forms, a nonphosphorylated IIA form and a phosphorylated IIO form, which are interconvertible and functionally distinct. *In vivo*, the pol IIO enzyme contains as many as 50 phosphorylated amino acids (primarily phosphoserine) within the CTD (30). During transcription initiation, pol IIA is recruited to the DNA template by the general transcription factors. The pol IIA CTD undergoes extensive phosphorylation and conversion to IIO during the transition from preinitiation complex to stable elongation complex. Several CTD kinase activities have been implicated in CTD hyperphosphorylation, each of which contains a cyclin and cyclin-dependent kinase subunit pair. The cdk7 and cyclin H subunits of the general transcription factor TFIIH catalyze phosphorylation of Ser-5 of the CTD heptapeptide (31). Other CTD kinases include the cdk8/cylin C pair found in the pol II holoenzyme (32, 33), CTDK-I, a heterotrimeric kinase with cdk-like and cyclin-like subunits (34), and P-TEFb, a regulator of polymerase elongation with a cdc2-like subunit (35, 36).

We recently showed that the recombinant *S. cerevisiae* and *S. pombe* guanylyltransferases Ceg1 and Pce1 bind specifically to the phosphorylated form of the CTD (2). Moreover, recombinant yeast cap methyltransferase Abd1 also binds specifically to CTD-PO4 (2). Phosphorylation at Ser-5 of the heptad repeat was necessary and sufficient to confer guanylyltransferase and methyltransferase binding capacity to the CTD (2). Here, we have extended this analysis to mammalian capping enzyme. The key finding is that the guanylyltransferase domain by itself binds to CTD-PO4, whereas the triphosphatase domain has no evident capacity to bind to the CTD. The CTD binding studies reported herein employed a CTD-PO4 ligand composed of recombinant GST-CTD-PO4 that was phosphorylated *in vitro*. Preliminary experiments show that the purified mouse guanylyltransferase domain Mce1(211–597) also binds to a chemically synthesized 42-amino acid phosphopeptide consisting of 6 tandem repeats of the CTD heptad sequence (YSPT**S**PS) in which all six Ser-5 residues are Ser-PO4.[2] The purified triphosphatase domain Mce1(1–210) does not bind to the 42-amino acid CTD phosphopeptide. These findings suggest that the mammalian RNA triphosphatase is targeted to the nascent pre-mRNA by virtue of its connection in *cis* to the guanylyltransferase.

Lower eukaryotes may have adopted an alternative strategy to achieve the same end. Although the guanylyltransferase and triphosphatase components of budding yeast are encoded separately, the Ceg1 and Cet1 proteins interact in *trans* to form a heteromeric enzyme complex that can be isolated from yeast extracts (28). Because Ceg1 by itself can bind CTD-PO4, it may well chaperone Cet1 to the pol II elongation complex. If this is the case, then our genetic complementation data raise the prospect that the mouse guanylyltransferase domain can engage to some degree in cross-species interaction with yeast triphosphatase. Alternatively, Cet1 may have its own capacity

[2] C. K. Ho and S. Shuman, unpublished data.

to bind the pol II transcription complex, be it through the CTD or some other constituent of the complex. In either case, it appears that poisoning the mouse triphosphatase active site in the full-length Mce1 enzyme antagonizes the yeast capping system *in vivo* in a way that deletion of the triphosphatase domain does not.

Recruitment of guanylyltransferase to the phosphorylated CTD neatly accounts for pol II specificity of capping and also provides a means of traffic control whereby CTD-interacting factors bind and dissociate from polymerase at the appropriate times in the transcription cycle without getting in each other's way. During preinitiation complex formation, the unphosphorylated CTD interacts with several general transcription factors and the SRB/mediator component of the pol II holoenzyme (37). Phosphorylation of the CTD presumably destabilizes these contacts and makes the CTD available for a novel set of interactions with the capping enzymes. Our findings contribute to an emerging picture of the CTD as a landing pad for macromolecular assemblies that regulate mRNA synthesis and processing (38, 39). Other recent studies indicate that protein components of the pre-mRNA splicing and 3′ cleavage-polyadenylation assemblies also bind to the CTD (40, 41). The role of CTD phosphorylation in those interactions remains to be clarified. Thus far, only the capping enzymes display a strict requirement for CTD phosphorylation.

REFERENCES

1. Shuman, S. (1995) *Prog. Nucleic Acids Res. Mol. Biol.* **50,** 101–129
2. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997) *Genes Dev.* **11,** 3306–3318
3. Cho, E., Takagi, T., Moore, C. R., and Buratowski, S. (1997) *Genes Dev.* **11,** 3319–3326
4. Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D., and Shatkin, A. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12898–12903
5. Shibagaki, Y., Itoh, N., Yamada, H., Nagata, S., and Mizumoto, K. (1992) *J. Biol. Chem.* **267,** 9521–9528
6. Shuman, S., Liu, Y., and Schwer, B. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 12046–12050
7. Yamada-Okabe, T., Shimmi, O., Doi, R., Mizumoto, K., Arisawa, M., and Yamada-Okabe, H. (1996) *Microbiology* **142,** 2515–2523
8. Shuman, S., and Hurwitz, J. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 187–191
9. Shuman, S., and Schwer, B. (1995) *Mol. Microbiol.* **17,** 405–410
10. Tsukamoto, T., Shibagaki, Y., Imajoh-Ohmi, S., Murakoshi, T., Suzuki, M., Nakamura, A., Gotoh, H., and Mizumoto, K. (1997) *Biochem. Biophys. Res. Commun.* **239,** 116–122
11. Yagi, Y., Mizumoto, K., and Kaziro, Y. (1983) *EMBO J.* **2,** 611–615
12. Yagi, Y., Mizumoto, K., and Kaziro, Y. (1984) *J. Biol. Chem.* **259,** 4695–4698
13. Takagi, T., Moore, C., Diehn, F., and Buratowski, S. (1997) *Cell* **89,** 867–873
14. Wang, S. P., Deng, L., Ho, C. K., and Shuman, S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 9573–9578
15. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77,** 51–59
16. Schwer, B., and Shuman, S. (1994) *Proc. Natl Acad. Sci. U. S. A.* **91,** 4328–4332
17. Shuman, S., Surks, M., Furneaux, H., and Hurwitz, J. (1980) *J. Biol. Chem.* **255,** 11588–11598
18. Shuman, S. (1982) *J. Biol. Chem.* **257,** 7237–7245
19. Denu, J. M., Stuckey, J. A., Saper, M. A., and Dixon, J. E. (1996) *Cell* **87,** 361–364
20. Ho, C. K., Van Etten, J. L., and Shuman, S. (1996) *J. Virol.* **70,** 6658–6664
21. Hakansson, K., Doherty, A. J., Shuman, S., and Wigley, D. B. (1997) *Cell* **89,** 545–553
22. Fresco, L. D., and Buratowski, S. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 6624–6628
23. Myette, J. R., and Niles, E. G. (1996) *J. Biol. Chem.* **271,** 11936–11944
24. Cong, P., and Shuman, S. (1995) *Mol. Cell. Biol.* **15,** 6222–6231
25. Myette, J. R., and Niles, E. G. (1996) *J. Biol. Chem.* **271,** 11945–11952
26. Yu, L., and Shuman, S. (1996) *J. Virol.* **70,** 6162–6168
27. Yu, L., Martins, A., Deng, L., and Shuman, S. (1997) *J. Virol.* **71,** 9837–9843
28. Itoh, N., Mizumoto, K., and Kaziro, Y. (1984) *J. Biol. Chem.* **259,** 13930–13936
29. Gunnery, S., and Mathews, M. B. (1995) *Mol. Cell. Biol.* **15,** 3597–3607
30. Dahmus, M. E. (1996) *J. Biol. Chem.* **271,** 19009–19012
31. Roy, R., Adamczewski, J. P., Seroz, T., Vermeulen, W., Tassan, J. P., Schaeffer, L., Nigg, E. A., Hoeijmakers, J., and Egly, J. M. (1994) *Cell* **79,** 1093–1101
32. Liao, S., Zhang, J., Jeffrey, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., van Vuuren, H. J. J., and Yound, R. A. (1995) *Nature* **374,** 193–196
33. Leclerc, V., Tassan, J. P., O'Farrell, P. H., Nigg, E. A., and Leopold, P. (1997) *Mol. Biol. Cell* **7,** 505–513
34. Sterner, D. E., Lee, J. M., Hardin, S. E., and Greenleaf, A. L. (1995) *Mol Cell. Biol.* **15,** 5716–5724
35. Zhu, Y., Pe'ery, T., Peng, J., Ramanathan, Y., Marshall, N., Marshall, T., Amendt, B., Mathews, M. B., and Price, D. H. (1997) *Genes Dev.* **11,** 2622–2632
36. Mancebo, H. S. Y., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Peng, J., Blau, C., Hazuda, D., Price, D., and Flores, O. (1997) *Genes Dev.* **11,** 2633–2644
37. Orphanides, G., Lagrange, T., and Reinberg, D. (1996) *Genes Dev.* **10,** 2657–2683
38. Steinmetz, E. J. (1997) *Cell* **89,** 491–494
39. Shuman, (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 12758–12760
40. Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R., Gentile, C., Gebara, M., and Corden, J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6975–6980
41. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G. H., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997) *Nature* **385,** 357–361