Molecular Cloning of CDK7-associated Human MAT1, a Cyclin-dependent Kinase-activating Kinase (CAK) Assembly Factor

Ann Yee, Michael A. Nichols, Lingtao Wu, Frederick L. Hall, Ryuji Kobayashi, and Yue Xiong

Department of Biochemistry and Biophysics [A. Y., Y. X.] and Curriculum in Molecular Biology and Genetics [M. A. N., Y. X.], University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280; Division of Orthopaedic Surgery, Research Institute of Children's Hospital Los Angeles, Los Angeles, California 90027 [L. W., F. L. H.]; Department of Molecular Pharmacology and Toxicology, University of Southern California School of Pharmacy, Los Angeles, California 90033 [F. L. H.]; and Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724 [R. K.]

Abstract

Mammalian CDK7 is a protein kinase identified as the catalytic subunit of cyclin-dependent kinase (CDK)-activating kinase and as an essential component of the transcription factor TFIIH that is involved in transcription initiation and DNA repair. We have identified in human cells a number of CDK7-associated cellular proteins that appear to fall into two classes based on their relative [35S] metabolic labeling intensity. One class of proteins present in CDK7 immunocomplexes as a minor fraction contains components of the TFIIH transcription complex such as p62 and p39\textsuperscript{ cyclin H}, whereas the other fraction contains four polypeptides (p35, p37\textsuperscript{ cyclin H}, p75, and p95) that are stoichiometrically associated with CDK7. Whereas the levels of association of p35, p37\textsuperscript{ cyclin H}, and p75 with CDK7 remain unchanged between density-arrested and proliferating Ewing sarcoma EW-1 cells, the association of p95 with CDK7 was significantly decreased as cells reached confluency. Through a large-scale immunopurification of CDK7 complexes and protein microsequencing, we have isolated a CDNA that encodes p35 and have shown that it is the human homologue of Mat1 that is involved in the assembly of CAK. MAT1 contains a highly conserved C\textsubscript{3}H\textsubscript{C} motif at its NH\textsubscript{2} terminus, a characteristic feature shared among RING finger proteins. The human MAT1 gene expresses a single 1.6-kb transcript, the steady-state level of which, like CDK7 and cyclin H, varies significantly in different cell lines and in different terminally differentiated tissues.

Introduction

CDKs provide the primary regulation for cell cycle progression, and there are at least three different mechanisms to regulate the enzymatic activity of a CDK: (a) cyclin binding; (b) subunit phosphorylation; and (c) association with a CDK inhibitor (see recent reviews in Refs. 1–4). Whereas binding to an appropriate cyclin is a necessary activation step, and association with CDK inhibitors is inactivating, phosphorylation on a CDK catalytic subunit can be either inhibitory or activating depending on the site of phosphorylation (reviewed in Refs. 2 and 3). Three conserved phosphorylation sites have been identified that correspond to Thr14, Tyr15, and Thr161 in human p34\textsuperscript{cdc2}. Phosphorylations of Thr14 by an as yet unidentified protein kinase and of Tyr15 by Weel/Mik1-related kinases are inhibitory. Both sites must be dephosphorylated by cdc25 protein tyrosine phosphatase to activate p34\textsuperscript{cdc2} and possibly other CDKs. On the other hand, phosphorylation of Thr161 is required for p34\textsuperscript{cdc2} activation. The enzyme responsible for this activation, termed CAK, was initially purified from Xenopus oocyte extracts based on its ability to phosphorylate \textit{in vitro} the Thr161 site of CDC2 (5). The catalytic subunit of CAK was then identified in starfish and Xenopus (6–8) and later in mammalian cells (9–11) as the MO15 gene product, initially isolated in a search for cdc2-related protein kinase in Xenopus (12). Phosphorylation of a CDK by CAK is blocked when the CDK subunit is bound to members of either the p21 or the p16 CDK inhibitor families (13, 14). Identification of a cyclin-related Mr 37,000 polypeptide associated with the catalytic subunit in mammalian cells, p37\textsuperscript{cyclin H}, and the reconstitution of CAK activity \textit{in vitro} when MO15 catalytic subunit was combined with p37\textsuperscript{cyclin H} demonstrated that CAK is itself a CDK (CDK7; Refs. 9, 15, and 16). The requirement of Thr161 phosphorylation for the activation of most, if not all, CDKs suggests a critical role of CDK7 in regulating the cell cycle.

CDK7 and cyclin \textsuperscript{H} were recently identified as essential components of the transcription factor TFIIH and implicated as the kinase that phosphorylates the COOH-terminal domain of the largest subunit of eukaryotic RNA polymerase II (17–20). In addition to its essential role in transcription initiation, TFIIH has also been demonstrated to be directly involved in DNA excision repair processes (21–24). These findings implicate CAK in diverse regulatory pathways important to physiological processes such as transcriptional control, DNA repair, and cell cycle regulation. In attempts to elucidate the mechanisms that regulate the function of CAK, we have undertaken a more detailed analysis of the subunit composition of the CDK7 complex in human cells. Our studies revealed the existence of two distinct populations of CDK7. This work further identified, in addition to the previously isolated p37\textsuperscript{cyclin H}, three novel CDK7-related associated cellular proteins (p35, p75, and p95). Molecular cloning of p35 identifies it as the human homologue of the mouse (25), Xenopus, and starfish \textit{MAT1} (26) that may be involved in the assembly of CAK complexes.

Materials and Methods

Cell Culture and Immunocytochemistry Procedures. HeLa (a human cervix epitheloid carcinoma cell line) and EW-1 (a human osteosarcoma cell line) were cultured in DMEM supplemented with 10% fetal bovine serum and antibiotics. Antibodies to CDK2 (27) and CDK7 (28) and procedures for \textsuperscript{[35S]}methionine metabolic labeling, immunoprecipitation, and SDS-PAGE have been described previously (27, 29).

Molecular Cloning of p35. We have previously reported a large scale immunopurification of CDK7 complexes from human HeLa cells using an anti-CDK7 antibody raised against a synthetic peptide corresponding to the COOH terminal 18 residues of human MO15/CDK7 (28). Approximately 2 \mu g of p35 were purified and subjected to an in-gel digestion with Achromobacter endopeptidase. High-performance liquid chromatography purification and peptide sequencing were as described previously (28, 30). Four peptide sequences were obtained (Ref. 28 and Fig. 2). Degenerate oligonucleotides were synthe-
sized based on the amino acid sequences derived from these peptides and were used as primers to amplify sequences from cDNA templates prepared from a HeLa cDNA library (Stratagene, Inc.) by PCR. Various PCR conditions were tried using different combinations of primers. Of the more than 100 subclones sequenced, amplification using 1 pair of primers derived from peptides K17 (5'-GAAATTCTAGA/AG/GA/AG/GCCN/CT/TTNTA/TG/CA/AG/GTA/TC)-CA-3', encoding EALELEYEY Q and containing an EcoRI site as indicated by the underline) and K6 (5'-AACTCAAGCTTACTCNCNCGNCNGCA/AG/G/AG/G/CT/CT/TTGT-3', encoding an antisense sequence for QDLAGGY and containing a XhoI site) resulted in the isolation of six independent clones, each containing a 138-bp fragment after 35 cycles of PCR amplification (denaturing at 94°C for 20 s, reannealing at 56°C for 20 s, and elongation at 72°C for 30 s). This sequence translates to a 46-amino acid residue open reading frame that also contains a proline residue preceding the K6 primer that was withheld from the oligonucleotide design for verification purposes. During the course of isolating the p35 cDNA, a mouse cDNA, p36 Mac(4) was reported that contained a region nearly identical to the 46-amino acid sequence amplified from human cells (amino acid residue 242–287; Ref. 25); further confirming the 46-amino acid sequence as human CDK7-associated p35. The 138-bp fragment was used as a probe to isolate the full-length human p35 by screening a HeLa cDNA library. All three cDNA clones isolated from this screen (H2, H7, and H8) were completely sequenced on an Automated Sequencing System (373A, Applied Biosystem).

Northern Analysis. For multiple tissue Northern analysis, 2 μg poly(A)+ RNA were isolated from different human tissues, resolved on a 1.2% agarose gel, and transferred to a nylon membrane (Clontech, Inc.). For Northern analysis in human cultured cell lines, 25 μg of total RNA were isolated from six different human cell lines, resolved on a 1.2% agarose gel, and transferred to a nylon membrane. A 1.3-kb EcoRI restriction fragment derived from cDNA clone H2 was labeled by the random priming labeling method and was used as a probe. The same blot was stripped and rehybridized with a 1.3-kb cDNA fragment corresponding to human MO15/CDK7 (10) and a 1.2-kb cDNA fragment corresponding to the human cyclin H (15, 16).

Results

To identify potential CDK7-associated cellular proteins, [35S]methionine-labeled lysates were prepared from a variety of human cell lines and immunoprecipitated with an antipeptide antibody specific to human MO15/CDK7. In addition to the previously identified p37 Cyclin H, this study revealed several polypeptides that appear to associate with CDK7 in a specific manner (Fig. 1). These candidate CDK7-associated proteins were effectively competed by the preincubation of the anti-CDK7 antibody with antigen peptide (Fig. 1) and did not show crossreactivity in immunoblots of total cell extracts with the anti-CDK7 antibody (data not shown). They include three polypeptides (p35, p75, and p95) that communoprecipitated with CDK7 at approximately the same intensity of [35S]labeling, whereas several additional proteins such as p62 and p89 seem to be present in the CDK7 complexes at much lower intensity. As determined from their gel mobility and by immunoblotting, both p62 and p89 appear to correspond to the p62 and p89ECC1 present in TFIIF transcription factor complexes (data not shown). Whether other faint protein bands (e.g., the two proteins migrating at Mr 48,000 and 70,000) that are present and compatible in the anti-CDK7 immunocomplexes also correspond to other components of the TFIIF complexes has not been determined due to the lack of suitable immunological reagents.

On the basis of their relative intensities, there appears to exist two populations of CDK7: (a) one complexing with four proteins (p35, p37 Cyclin H, p75, and p95), each of which appears to associate with CDK7 stoichiometrically; and (b) the other containing several polypeptides that complex with CDK7 to a lesser degree. The second subpopulation includes at least two proteins (p62 and p89ECC1) that are essential components of the TFIIF complex and may represent the TFIIF-bound fraction of CDK7. Because CDK7 and cyclin H are present in TFIIF complexes stoichiometrically to other TFIIF components4, the presence of p62 and p89ECC1 (and potentially other TFIIF subunits as well) at a much lower level in CAK/CDK7 complexes suggests that the major fraction of CDK7 is not bound to TFIIF. This TFIIF-unbound fraction of CDK7 appears to mainly associate in vivo with four intensely [35S]labeled proteins that are present in CDK7 immunocomplexes at a level similar to that of CDK7 and at a much higher level than that of p62 and p89ECC1. It remains to be determined whether CDK7 associates with these proteins as multiple binary, ternary, quaternary complexes, or as a single complex consisting of all five proteins. Upon reaching confluency, growth of EW-1 cells is arrested, and synthesis and/or association of S-phase cyclin A with CDK2 is decreased (Fig. 1, Lanes 3 and 7). We noticed that the level of CDK7-associated p95, but not other CDK7-associated proteins, was significantly lower in density-arrested EW-1 cells than in log-phase growing cells, as determined by the metabolic labeling (Fig. 1). Whether this is due to decreased level of p95 mRNA, protein, or its association with CDK7, and whether this protein may play a role in regulating the function of CAK, must await its molecular cloning.

We isolated a cDNA from HeLa cells that encodes one of three uncharacterized, major CDK7 associated polypeptides (p35) after a large-scale anti-CDK7 immunoprecipitation and protein microsequencing (Ref. 28; see also “Materials and Methods”). Three cDNA

4 Z-Q. Pan and A. Sancar, personal communication.
clones (H2, H7, and H8) obtained from this screening were completely sequenced, and all contain sequences highly related to the MAT1 sequence that was very recently isolated from mouse, Xenopus, and starfish (25, 26). Following proposed nomenclature, we therefore refer to the gene encoding human CDK7-associated p35 also as Mat1 (for ménage à trois 1). The H8 clone contains a poly(A) tail (42 A residue) and about 500 bp of MAT1-related sequence preceding the poly(A) tail, but 5' sequence upstream from the middle of the coding region completely diverged from mouse, Xenopus, and starfish Mat1, as well as the other two human cDNA clones (corresponding to the nucleotide 407 of H2 as presented in Fig. 2A). The upstream divergent sequence in H8 is nearly identical to a human mitochondria sequence (GenBank accession no. D53941), suggesting that it might have been fused with p35 during cDNA library construction. The H7 clone contains a 1.2-kb insert and is highly related to MAT1. This clone, however, is truncated at its 5' end and is missing the first 106 amino acid residues compared to MAT1 in other species or the human cDNA clone H2 (Fig. 2A). The H2 clone contains a 4.4-kb insert consisting of four EcoRI restriction fragments, of which one 1.3-kb EcoRI fragment contains all four of the peptide sequences obtained from protein microsequencing (underlined in Fig. 2A), confirming it as human CDK7-associated p35. The remaining 3.1-kb sequence is not related to MAT1 nor to any other sequence currently deposited in database, and its origin has not been determined. Comparison of the sequence in the cDNA clone H2 with either clone H7 or mouse p36Mat1, however, revealed a single base pair (nucleotide A) deletion at nucleotide position 540, which results in a reading frame shift and a truncated protein of Mr 20,000 (indicated by a hyphen in Fig. 2A). To verify this alteration, we produced a protein encoded by the 1.3-kb EcoRI fragment by in vitro translation and by expression in bacteria and confirmed that it indeed encodes a M, 20,000 protein (data not shown). The truncated M, 20,000 protein lost nearly half of its carboxy terminal region, but still retained the RING finger portion located in its amino terminal region (see below). Whether this alteration originated in the cDNA library construction or represents an in vivo mutation of Mat1 in HeLa cells has not been determined. Nor have we determined whether the truncated M, 20,000 protein can still interact with CDK7-cyclin H or interfere with the interaction of wild-type p35Mat1 with CDK7-cyclin H.

Conceptual translation of human Mat1 as compiled from the cDNA clones H2 and H7 revealed a 309-amino acid reading frame that encodes a M, 35,800 protein. Human Mat1 is very closely related to mouse Mat1: 87% identity at nucleotide level and 95% identity between the two proteins. Human MAT1 also shares 79 and 51% protein sequence identity with Xenopus and starfish Mat1, respectively. Sequence conservation between MAT1 from different species is significantly higher at the NH2-terminal portion. This portion of MAT1 contains a C2HC4 motif presented in many RING finger proteins (Ref. 31; Fig. 2B). Within the RING finger region, human MAT1 is identical to mouse MAT1 and nearly identical to Xenopus and starfish MAT1, suggesting an essential role of this domain for the function of MAT1. Although the zinc-binding RING finger motif has been identified in a number of proteins involved in diverse biological processes (31), neither its biochemical properties nor its physiological function(s) are well understood. We have not determined whether this motif is necessary for the association of MAT1 with CDK7 or CDK7-cyclin H.

Using the 1.3-kb EcoRI fragment derived from cDNA clone H2 that contains p35Mat1 as a probe (Fig. 2A), we carried out a Northern blot analysis to determine the expression of p35Mat1 mRNA in different human tissues and cultured cell lines. Under high stringency conditions, this probe detected a single band of approximately 1.5 kb (Fig. 3, A and B). The same multiple tissue Northern blot was stripped and rehybridized with a probe corresponding to human MO15/CDK7 and cyclin H, respectively (Fig. 3A). The steady-state level of Mat1 mRNA, like that of both MO15/CDK7 and cyclin H, varied significantly between different human tissues. The highest level of Mat1 was observed in colon and testis; moderate levels were present in thymus, prostate, ovary, and small intestine. Mat1 mRNA is present in both leukocyte and spleen at a nearly undetectable low level. It was of interest that we noticed that the pattern of mRNA expression/accumulation of both CAK/CDK7 and cyclin H are somewhat similar to that of Mat1. Variations in the level of Mat1 mRNA were also found in cultured human cells (Fig. 3B). Consistent with the high level of p35Mat1 protein initially observed in EW-1 cells (Fig. 1), Mat1 mRNA was found to be high in EW-1 cells (Fig. 3B).

Discussion

CAK was initially partially purified from Xenopus eggs through its activity in phosphorylating the threonine 161 on CDC2 in vitro (5). Further purification from both Xenopus eggs and starfish oocytes led to the discovery that the MO15 gene, a previously identified cdc2-related protein kinase in Xenopus, encodes the catalytic subunit of this activity [p40MO15, Refs. 7, 8, 12]. Copurification of additional polypeptides with p40MO15 and the lack of independent CAK activity in bacteria-produced p40MO15 prompted the search for and subsequent identification in mammalian cells of a cyclin-related protein (cyclin H) as the first regulatory subunit physically associated with p40MO15 (which led to the renaming of MO15 as CDK7; Refs. 9, 15, 16). The final enzyme fraction of CAK from the initial biochemical purification contained only two major polypeptides (cyclin H and the catalytic subunit MO15/CDK7 (6, 15)). Immunopurification using antibodies specific to the catalytic subunit, however, revealed additional cellular proteins present in anti-CDK7 immunocomplexes (Refs. 9, 28; Fig. 1). Comparing the relative intensity of these CAK-related-associated proteins with CDK7, there appears to exist at least two distinct populations of CDK7: TFIIH bound and TFIIH unbound (Fig. 1). The fraction of CDK7 that contains several proteins that are present in anti-CDK7 immunocomplexes at a level much lower than that of CDK7 appear to represent a TFIIH-bound fraction of CDK7 because at least two of these faint CDK7-associated proteins (p62 and p89SRC3) correspond to the components of TFIIH. The other fraction of CDK7 immunocomplexes includes four polypeptides (p35, p37, p75, and p95) that are present at a level similar to that of CDK7 and at a much higher level than that of p62 and p89SRC3.

Molecular cloning of Mat1 and the existence of two additional CDK7-associated proteins (p75 and p95) impose a new level of regulation on the conventional binary CDK7-cyclin H complex. MAT1 is a new member of the RING finger protein family with a characteristic C2HC4 motif and acts as an assembly factor to stabilize the binary CDK7-cyclin H association in the absence of phosphorylation of a Thr 161 equivalent residue in CDK7 [Thr 176 in Xenopus CDK7 and Thr 170 in mouse CDK7 (25, 26)]. Analogous Thr 161 phosphorylation of CDK7 by an as yet unidentified kinase (CASKAK) releases this requirement and allows CDK7-cyclin H to form heterodimer (25). Together with the finding that CDK7 may further associate with additional cellular proteins (Fig. 1), these results indicate that there may exist in vivo multiple pathways of CDK7-cyclin H (CAK) assembly and activation.

The original model of a cell cycle kinase consisted of a catalytic subunit (CDK) and a regulatory subunit, (cyclin). The discovery of a group of inhibitory proteins (CDK inhibitors) provided the first evidence of the regulation of CDK activity by the physical complexing of CDKs with other cellular proteins. Recent identification of several novel CDK-associated proteins has revealed an additional regulatory pathway that regulates the activity of cell cycle kinases (Refs. 25, 26, and 32 and this work). Rather than directly activating or inhibiting
Fig. 2. cDNA and predicted amino acid sequence of human MAT1. A, nucleotide and amino acid sequences of human MAT1. Numbers for both nucleotides and amino acids are given, and the stop codon is indicated by an asterisk. A hyphen at position 540 indicates a 1-bp deletion in cDNA clone H2 as compared to the cDNA clone H1 and mouse MAT1. Nucleotide sequences in lower case print preceding the 5' end and following the 3'-end are two EcoRI sites used in cloning. Mouse (MAT1-Mm), Xenopus (MAT1-X1), and starfish (MAT1-Mg) sequences were included underneath human sequence as a comparison. Amino acid residues identical to human MAT1 are represented by a hyphen. The microsequenced peptide fragment is underlined. B, protein sequence comparison of the RING finger motif. RING finger domain of MAT1 proteins from four different species are aligned with the RING finger domain from four representative RING finger-containing proteins (RAG1-Hs (33), XNF7-X1 (34), COP1-At (35), and RING1-Hs (36)).

enzymatic activity upon their association with CDK, these newly discovered proteins control the assembly of CDKs into higher order complexes (e.g., MAT1 in assembling CDK7-cyclin H (25, 26) and SKP2 in assembling CDK2-cyclin A (32)). The challenging question that remains to be answered is how the activity of CDK is regulated by these assembly factors: (a) by regulating the timing or specificity of cyclin association and other complex assemblies and interactions; (b) by targeting the CDKs to different substrates; (c) by changing the subcellular localization of the CDKs, or (d) by preventing the binding of and thus inhibition by CDK inhibitors.

6061
cultured cell lines as indicated at the top of each lane were hybridized with a 1.3-kb probe and a cyclin H probe. B. twenty-five ug of total RNA prepared from six different human tissues as indicated at the top of each lane were hybridized with a 1.3-kb probe derived from human Mat1 cDNA. A 1.5-kb discrete band was detected by this probe different human tissues as indicated at the top of each lane were hybridized with a 1.3-kb antibodies, Kim MacFarland for excellent technical support, Yan Li for help with Northern blotting, David Franklin for assistance with photography, and 5. Solomon, M. J., Lee, T., and Kirschner, M. w. Role of phosphorylation in p34\(^{\text{cdk2}}\)
Molecular Cloning of CDK7-associated Human MAT1, a Cyclin-dependent Kinase-activating Kinase (CAK) Assembly Factor

Ann Yee, Michael A. Nichols, Lingtao Wu, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/55/24/6058

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.