Identification of a developmentally regulated protein-tyrosine kinase by using anti-phosphotyrosine antibodies to screen a cDNA expression library

(embryonic development)

ELENA B. PASQUALE AND S. J. SINGER

Department of Biology, B-022, University of California, San Diego, La Jolla, CA 92039

Contributed by S. J. Singer, April 24, 1989

ABSTRACT To identify the protein-tyrosine kinases that are expressed during chicken embryonic development, a 10-day chicken embryo cDNA expression library was screened with anti-phosphotyrosine antibodies. Of the positive clones isolated, many encoded the same protein-tyrosine kinase, which we designate Cek1 (chicken embryony kinase 1). Its amino acid sequence suggests that the Cek1 protein is a transmembrane tyrosine kinase and presumably the receptor for an unknown ligand. Antibodies prepared to the cloned Cek1 kinase recognize, in immunoblotting experiments, two protein bands with apparent molecular weights of 100,000 and 110,000. The Cek1 protein was detected in many chicken embryonic tissues, but not in the corresponding adult tissues (with the exception of brain). The Cek1 kinase appears to be very closely related to two protein-tyrosine kinases whose partial sequences have been recently determined, human Fig and mouse Bek. Cloning using anti-phosphotyrosine antibodies has allowed us to detect, in addition to Cek1, several other protein-tyrosine kinases that are expressed during chicken embryonic development, some of which have not been previously identified.

Protein-tyrosine kinases are thought to play a role in the control of cellular growth and differentiation. This assumption is based on the observation that a number of transforming retroviruses carry homologues of cellular genes encoding protein-tyrosine kinases and that several growth factor receptors display tyrosine kinase activity (for reviews see refs. 1 and 2). Various embryonic tissues have been examined for the presence of protein and/or mRNA corresponding to protooncogenes encoding protein-tyrosine kinases and to growth factor receptors with tyrosine kinase activity (for reviews see refs. 3 and 4). Despite evidence for the presence of several tyrosine kinases during development, the demonstration that phosphorylation of proteins on tyrosine occurs at substantial levels during certain stages of embryonic development is only recent (5). By using anti-phosphotyrosine (anti-P-Tyr) antibodies, proteins phosphorylated on tyrosine were found to be present in all the embryonic tissues examined, including heart, thigh, gizzard, intestine, lung, kidney, brain, and lens. The overall level of protein tyrosine phosphorylation in most of the tissues decreased during later embryonic development and was very low or undetectable in the tissues of the adult. These results indicate that a number of protein-tyrosine kinases must be expressed in different tissues and be active during embryonic development.

To develop a systematic approach for the identification of the protein-tyrosine kinases that are expressed during chicken embryonic development (which could include already known as well as previously unknown tyrosine kinases), we decided to use anti-P-Tyr antibodies for the screening of embryonic cDNA expression libraries. This approach was based on the observation that the protein-tyrosine kinase encoded by the v-abl oncogene of Abelson leukemia virus could be expressed in its enzymatically active form in Escherichia coli (6). Even a 1.2-kilobase (kb) fragment of v-abl, comprising little more than the catalytic domain, was shown to encode an active tyrosine kinase in E. coli (7). Further, a number of cellular tyrosine kinases are active after truncation of portions of their amino-terminal, noncatalytic domain (8, 9). Therefore it seemed likely that many of the protein-tyrosine kinase fragments encoded by cDNA inserts may be expressed in bacteria in an active form. Since protein-tyrosine kinases are capable of autophosphorylation and, under appropriate conditions, of phosphorylating nonphysiological substrates, the phosphorylation on tyrosine of most recombinant kinases as well as that of bacterial proteins was thought likely to occur. In addition, since in bacteria endogenous tyrosine kinase activity is undetectable (refs. 10 and unpublished data), only the clones expressing active tyrosine kinases were expected to be positively identified by anti-P-Tyr antibodies.

 Accordingly, cDNAs encoding portions of what we have demonstrated to be several different protein-tyrosine kinases were isolated with anti-P-Tyr antibodies, after screening of a 10-day chicken embryo library. Since many of the clones encoded portions of the same protein-tyrosine kinase, our initial experiments were directed to its characterization. The longest cDNA insert encoding this kinase was sequenced and appeared to contain the entire coding sequence of a protein with the features of a transmembrane tyrosine kinase receptor. The cloned kinase was detected, with specific polyclonal antibodies, in many chicken embryonic tissues, but not in most of the corresponding adult tissues. This developmentally regulated protein-tyrosine kinase, with apparent Mr = 100,000, is provisionally referred to as Cek1 (chicken embryo kinase 1; gene symbol, cek1).*

While these experiments were in progress (11), we learned that others were also using anti-P-Tyr antibodies to screen cDNA libraries (12-14).

MATERIALS AND METHODS

Cloning Using Anti-P-Tyr Antibodies. Polyclonal anti-P-Tyr antibodies were prepared by immunizing rabbits with the insoluble material obtained after sonication of bacteria expressing the v-abl-encoded tyrosine kinase in its active form (5, 6). The anti-P-Tyr antibodies from the immune serum were affinity-purified as described (5). Protein-

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M24637).
tyrosine kinases were cloned by screening a 10-day chicken embryo recombinant cDNA library (Clontech) constructed in phage Agt11 and expressed in E. coli y1090 according to the method of Young and Davis (15). Positive clones producing recombinant β-galactosidase–kinase fusion proteins were identified by probing duplicate nitrocellulose filters with the affinity-purified anti-P-Tyr antibodies (2 μg/ml) preabsorbed on E. coli proteins, followed by 125I-labeled protein A (0.25 μCi/ml; ICN; 1 μCi = 37 kBq), as described for immunoblotting (5).

**Sequencing.** Each cDNA insert isolated with anti-P-Tyr antibodies was subcloned in the phagemid pBluescript (Stratagene) and mapped with several restriction enzymes, and portions were sequenced by the chain-termination method (16) with double-stranded plasmid DNA. The complete sequence of the longest of the cek1 cDNA inserts (Fig. 1a), was obtained after subcloning in pBluescript the shorter cek1 inserts isolated as well as short portions of the cek1 cDNA, obtained by using suitable restriction enzymes. The noncoding strand complementary to nucleotides 1–150 in Fig. 1a was sequenced by using the oligodeoxynucleotide 5'-GGCCAGTGGGACTCCAC-3' (purchased from The Midland Certified Reagent Company, Midland, TX) as a primer, since suitable restriction sites were not present. The DNA sequence obtained was analyzed with the program ANALYSEQ (18), and the deduced amino acid sequences were compared with the sequences in the National Biomedical Research Foundation protein sequence data base (release of June 30, 1988) by using the program FASTA (19).

**Preparation of Antibodies Specific for the Cloned Tyrosine Kinases.** One of the longest cek1 cDNA inserts (2500 base pairs) was subcloned in the expression vector pEX (Boehringer Mannheim; ref. 20) to obtain high levels of expression of the β-galactosidase–Cek1 fusion protein. The insoluble

**FIG. 1.** (Figure continues on the opposite page.)
RESULTS

To identify the protein-tyrosine kinases that are expressed during chicken embryonic development, a 10-day chicken embryo cDNA expression library constructed in phage agt11 was screened with anti-P-Tyr antibodies. Of the 250,000 clones examined, 24 tested positive with the anti-P-Tyr antibodies. These positive clones encoded protein-tyrosine kinase fragments that were active in E. coli, since they were detected with anti-P-Tyr antibodies, even though their kinase domains were fused to the carboxyl-terminal end of the Agt11-encoded β-galactosidase. All the positive clones contained cDNA inserts of at least 1200 base pairs. The β-galactosidase-tyrosine kinase fusion proteins were usually the proteins most heavily phosphorylated on tyrosine in bacterial cell extracts, but, in some instances, several bacterial proteins also appeared to be phosphorylated on tyrosine in the positive clones (data not shown).

To characterize the cDNAs isolated with anti-P-Tyr antibodies, a number of the inserts were subcloned into the phagemid pBluescript for analysis with restriction enzymes and sequencing. The restriction maps of 10 of the inserts indicated that they represented different portions of the same gene (data not shown). Since the protein-tyrosine kinase encoded by this gene was likely to be abundant in a 10-day chicken embryo and, therefore, to play a prominent role during chicken embryonic development, we decided to focus our initial efforts on the characterization of this tyrosine kinase, designated Cek1 (chicken embryo kinase 1).

The complete nucleotide sequence of the longest cek1 cDNA insert isolated was determined and found to extend for 2965 nucleotides and to include a poly(A) sequence at the 3' end (Fig. 1a). Analysis of the cek1 cDNA sequence showed that an open reading frame of 819 bases (corresponding to a calculated molecular weight of 92,000) is flanked by 55 nucleotides of 5' untranslated sequence and 453 nucleotides of 3' untranslated sequence. We think that the ATG codon located at position 56 represents the initiation methionine, since an in-frame termination codon is located 36 bases

| b       | cek1    |     |    | flg    |     | 1     |     |    | bek    |     | 1     |     |    | Cek1  |     | 1     |     |    | Flg   |     |    | Bek   |     |    |
|---------|---------|-----|----|-------|-----|-------|-----|----|--------|-----|-------|-----|----|-------|-----|-----|--------|-----|----|--------|-----|----|
| cek1    | 199     | 451 | 351| flg    | 1   |       |     |    | bek    | 1   |       |     |    | cek1  |     | 1    |     |    | bek    |     |    | Cek1  |     | 1   |
| flg     | 249     | 499 | 251| bek    | 1   |       |     |    | bek    | 28  | KEAVD |     |    | flg   | 231 | 51  |     |    | bek    |     |    | Cek1  |     | 213 |
| cek1    | 299     | 549 | 399| flg    | 151 |       |     |    | flg    | 201 | .G...H...M...A...L       |
| cek1    | 349     | 499 | 301| flg    | 251 |       |     |    | flg    | 201 | .G...H...M...A...L       |
| cek1    | 399     | 549 | 399| flg    | 151 |       |     |    | flg    | 201 | .G...H...M...A...L       |

FIG. 1. (a) Nucleotide sequence and predicted amino acid sequence of the longest cek1 cDNA clone isolated. The presumed signal peptide sequence is underlined, the possible sites of N-glycosylation are underlined with a dotted line, the putative transmembrane domain is indicated by a solid bar, the ATP binding site is indicated by asterisks, and the tyrosine that is the putative site of autophosphorylation is indicated by a triangle. Nucleotides and amino acids are numbered at right. The nucleotides are numbered starting from the first nucleotide following the EcoRI site of agt11. (b) Comparison of the published human Flg and mouse Bek partial amino acid sequences with the corresponding portion of the Cek1 sequence. The three sequences are aligned. Dots replace residues in Fig and Bek that are identical to the corresponding residue in Cek1. Hyphens represent gaps introduced in the sequences to maximize alignment. Amino acids are numbered at left according to a (Cek1), ref. 17 (Fig), and ref. 12 (Bek).
upstream (Fig. 1). Despite the in-frame termination codon preceding the initial methionine, the Agt11 clone containing the cDNA insert shown in Fig. 1a must produce Cek1 protein, since it was isolated by using anti-P-Tyr antibodies to detect the tyrosine kinase activity of Cek1 produced in E. coli (see refs. 14 and 15 and references therein for examples of insert sequences that would not be expected to be expressed when carried in Agt11 but were nevertheless isolated by using antibodies to detect their protein products).

The putative initiation methionine is followed by a stretch of hydrophobic amino acids that likely represents a signal peptide sequence for translocation into the endoplasmic reticulum. If the putative signal sequence is subsequently cleaved, this most probably involves proteolysis between residues 21 and 22 (21). In addition to the signal sequence, the deduced amino acid sequence of Cek1 contains features characteristic of cell surface glycoprotein receptors. These include a stretch of hydrophobic amino acids, residues 375–395 in Fig. 1a, with the properties of a transmembrane domain. This putative transmembrane region is followed by Lys-396 and Lys-398, presumably representing a membrane stop-transfer signal. Three consensus sequences for asparagine-linked glycosylation (underscored with dots in Fig. 1a) are present in the presumed extracellular region, amino-terminal to the putative transmembrane region. The region carboxyl-terminal to the putative transmembrane domain, which is likely to represent the cytoplasmic catalytic domain, contains stretches of amino acids typical of tyrosine kinases (2).

After comparison of the deduced amino acid sequence of Cek1 with the sequences stored in a protein sequence data base (see Materials and Methods), it appeared that cek1 cDNA represented a protein-tyrosine kinase that had not been previously sequenced. However, inspection of tyrosine kinase partial sequences published most recently showed that the cek1 gene encodes a chicken tyrosine kinase closely homologous to a human putative tyrosine kinase designated Flg (17) (93% identity) and to a mouse tyrosine kinase designated Bek (12) (85% identity in the catalytic domain). In Fig. 1b the published deduced partial amino acid sequences of the Fig and Bek kinases are compared with the sequence of the Cek1 kinase. The three protein-tyrosine kinases may represent either the same protein in different species or very closely related proteins.

To identify the protein encoded by the cek1 gene and study its expression during embryonic development, anti-Cek1 antibodies were prepared by using a β-galactosidase-Cek1 fusion protein as the immunogen. After affinity purification, the anti-Cek1 antibodies were used in immunoblotting experiments to identify the protein in chicken tissue extracts and cultured cells. The anti-Cek1 antibodies labeled, to different extents, only one or two proteins with apparent Mₐ, 100,000 and 110,000 in extracts of several different tissues from 10-day embryos (Fig. 2a), but not in tissues from adult chicken (with the exception of brain, which was faintly labeled; Fig. 2b). The anti-Cek1 antibodies also labeled a Mₐ, 100,000 protein in cultured chicken embryo fibroblasts and, to the same extent, a Mₐ, 100,000 protein in Rous sarcoma virus-transformed chicken embryo fibroblasts (Fig. 2a). In immunoblotting control experiments (data not shown), the anti-Cek1 antibodies eluted from the Mₐ, 100,000–110,000

![Fig. 2. Immunoblotting of 10-day embryonic (a) and adult (b) chicken tissue extracts with polyclonal antibodies specific for the Cek1 protein-tyrosine kinase followed by 125I-labeled protein A. (a) Lanes: 1, blood; 2, intestine; 3, liver; 4, kidney; 5, lung; 6, brain; 7, gizzard; 8, heart; 9, thigh; 10, Rous sarcoma virus-transformed chicken embryo fibroblasts; 11, chicken embryo fibroblasts. (b) Lanes: 1, intestine; 2, liver; 3, kidney; 4, lung; 5, brain; 6, gizzard; 7, heart; 8, thigh. Arrow indicates the position of Cek1 in adult brain extract. Molecular weight standards used were myosin heavy chain (Mₐ, 200,000), β-galactosidase (Mₐ, 120,000), α-actinin (Mₐ, 100,000), bovine serum albumin (Mₐ, 68,000), and actin (Mₐ, 43,000).

region of an immunoblot of an extract of 10-day brain labeled the β-galactosidase-Cek1 fusion protein, confirming that the antibodies recognizing the Lanes 100,000–110,000 proteins in tissues and cultured cells were those specific for the cloned Cek1 protein-tyrosine kinase.

**DISCUSSION**

The large number of cDNA clones encoding protein-tyrosine kinases that we have isolated with anti-P-Tyr antibodies suggests that a surprisingly large fraction (we calculate about 0.1%) of the mRNA from a 10-day chicken embryo encodes proteins with tyrosine kinase activity and that, therefore, protein tyrosine phosphorylation has important functions during developmental processes. The expression of protein-tyrosine kinases may be especially high during certain stages of embryonic development, in agreement with the extensive protein tyrosine phosphorylation that was shown to occur in many tissues from chicken embryos 7–10 days old (5).

Of the 24 cDNA clones that we isolated by using anti-P-Tyr antibodies, 10 encode a protein-tyrosine kinase, cek1, that appears to be widely distributed in chicken embryonic tissues (Fig. 2a). Low levels of Cek1 were detected in adult chicken brain, but not in any other adult chicken tissues (Fig. 2b), indicating that the expression of cek1 is developmentally regulated and presumably not essential for cell survival. That the level of Cek1 is not elevated in Rous sarcoma virus-transformed chicken embryo fibroblasts compared to chicken embryo fibroblasts (Fig. 2a) suggests that its abundance does not simply correlate with rapid growth.

Cek1 appears likely to be a transmembrane receptor-like kinase, since it contains a potential signal peptide and a putative transmembrane region. If the potential signal peptide functions to initiate translocation of the kinase into the
endoplasmic reticulum and is subsequently cleaved, the mature protein would have an amino-terminal domain exterior to the plasma membrane, a single transmembrane domain, and the catalytic domain protruding from the cytoplasmic face of the plasma membrane. The molecular weight of 89,000 calculated from the cek1 cDNA sequence (not including the putative signal sequence) is in reasonable agreement with the apparent molecular weight of about 100,000 determined for the Cek1 protein by SDS/PAGE, especially when one considers that the Cek1 kinase may be glycosylated. The extracellular domain may serve as a receptor binding site for a ligand, such as a growth factor, as in the cases of the epidermal growth factor receptor and the platelet-derived growth factor (PDGF) receptor (22, 23). Cek1 may therefore be activated in developmental processes by interacting, through its extracellular region, with a specific soluble ligand or with a ligand present on the surface of neighboring cells (24). If so, that ligand is at present unknown.

Although the sequence of the catalytic domain of the Cek1 kinase is most similar to the corresponding sequence of the PDGF receptor and the c-fms and c-kit gene products (2), its extracellular domain contains an unusual stretch of acidic amino acids (residues 125–132) and evenly distributed cysteine residues. The 87-amino acid domain located between the putative transmembrane region and the nucleotide binding site is larger than the corresponding domains (of about 50 amino acids) of most known tyrosine kinase receptors. The inserted sequence that interrupts the kinase domain is very short in comparison with those of the PDGF receptor, c-Fms, and c-KIT and more similar in size to that of the insulin receptor (2). All of these features suggest that Cek1 may be a member of a previously uncharacterized transmembrane receptor tyrosine kinase family.

However, two tyrosine kinase genes that appear very closely related to cek1 have been recently isolated and partially sequenced: flg (21) and bek (12). Flg (for fms-like gene) was isolated from a human endothelial cell cDNA library by using the v-fms oncogene as a probe. Flg mRNA (of 4.5 kb) was detected by RNA gel blot analysis in human fibroblasts and placenta as well as in rat heart, eye, brain, kidney, lung, and testes. The expression of the flg gene was also found to be modulated during endothelial cell differentiation. Flg was thought to lack a transmembrane domain in close proximity to the kinase domain (17). According to our analysis, however, the sequences of both Cek1 and Flg (Fig. 1) contain a similar stretch of hydrophobic amino acids with the features of a transmembrane domain. The bek (for bacterially expressed kinase) gene was identified with anti-P-Tyr antibodies in an adult mouse liver cDNA library. One bek insert, extending from the beginning of the catalytic domain to the poly(A) tail, was isolated after screening 500,000 clones. By RNA gel blot analysis, bek transcripts (of 4.3 kb) were found in RNA from adult mouse liver, lung, brain, and kidney, but not heart and spleen.

Although the partial sequences of Flg and Bek are very similar to portions of the complete sequence of Cek1 (Fig. 1b), it is not known whether the three proteins represent the same gene (and the differences in their DNA and amino acid sequences are the result of evolutionary changes) or they are different but very closely related genes. The fact that the Cek1 protein is not detectable in most adult tissues, whereas flg and bek mRNAs were detected in several adult tissues, may indicate that the cek1 gene is different from the flg and bek genes or may simply reflect a greater sensitivity of RNA gel blotting compared with immunoblotting techniques.

Cloning using anti-P-Tyr antibodies for rapid screening and detection appears to be a method of general application for the systematic isolation of cellular protein-tyrosine kinases and for confirmation of their catalytic activity. The method is suitable for the identification of novel tyrosine kinases and is therefore particularly useful for the detection of protein-tyrosine kinases involved in developmental processes, only a fraction of which may be known at this time. In fact, after screening about 25% of a 10-day chicken embryo cDNA library, we isolated, in addition to the 10 cek1 clones described, a tyrosine kinase gene (which we have designated cek1) very closely homologous to the cek1 gene but distinct from it. Four additional clones were found to encode portions of p60<sup>tyr</sup>, a known intracellular tyrosine kinase very similar to p60<sup>tyr</sup>-src (25). Interestingly, the expression of c-yes mRNA in chicken tissues has been reported to be much higher than that of other protooncogenes from the avian sarcoma family (26). The deduced partial amino acid sequences of several of the other clones isolated did not give any indication regarding their identity, suggesting that they may encode still other protein-tyrosine kinases that have not been previously identified. Since some of the cloned tyrosine kinases are encoded by only one or two of the cDNAs isolated, it is likely that the screening of the remaining 75% of the library will lead to the identification of additional tyrosine kinases involved in embryonic development.

We thank P. A. Maher and R. A. Lindberg for helpful discussions and D. J. G. Rees for teaching one of us most of the molecular biology techniques used in this project. This work was supported by U. S. Public Health Service Grant GM-15971 (S.J.S.) and by an Arthritis Investigator Award (E.B.P.). S. J. S. is an American Cancer Society Research Professor.