ABSTRACT We have identified protein kinase genes of Dictyostelium by using highly conserved amino acid sequence motifs to design the synthesis and amplification of DNA fragments by polymerase chain reactions (PCRs). Cloning and sequencing the PCR products have revealed five different members of the protein kinase multigene family. These five putative kinases showed varying degrees of amino acid sequence similarity (40–70%) to protein kinases in data bases and contained invariant amino acid residues characteristic of protein kinases. DNA from PCR was labeled and used to isolate several Agt11 cDNA clones, including one full-length one (Dd kinase-2). The nucleotide sequence of Dd kinase-2 contained a region identical to one of the cloned kinase fragments amplified by PCR, and based on the deduced amino acid sequence Dd kinase-2 encodes a protein of 479 amino acids. A 350-amino acid kinase domain at the C-terminal end shows high homology to the catalytic domains of protein kinase A, protein kinase C, S-6 kinase of Xenopus, and the suppressor of cdc25 of yeast. The N-terminal domain is highly basic and also contains alternating threonine/proline residues. The cDNA hybridized to a single copy gene but to two differentially regulated mRNAs—a 2.0-kilobase mRNA that is expressed in vegetative cells and a 2.2-kilobase mRNA that is expressed during development. The larger mRNA is induced by cAMP by using a cell-surface receptor-mediated signal transduction pathway.

Reversible protein phosphorylation plays an essential role in signal transduction in both prokaryotes and eukaryotes (for reviews, see refs. 1 and 2). Using conserved amino acid sequence domains in known protein kinases (3), a large number of putative protein kinase genes have been cloned by hybridization (4) to DNA or oligonucleotide fragments (5, 6) and more recently by polymerase chain reaction (PCR) (7). The function of most protein kinases will be easier to determine in genetically tractable organisms, such as Dictyostelium discoideum, in which antisense inactivation of RNA and gene disruption are possible and in which molecules involved in signal transduction reactions have already been identified. During starvation, extracellular cAMP, a chemoattractant, binds to cell-surface receptors and causes unicellular amebae to aggregate (8). Extracellular cAMP also regulates the expression of several cell-type-specific genes by binding to cell-surface cAMP receptors (9–13). Thus, both receptor-mediated chemotaxis and gene induction are signal transduction events. Many of the components of the signal transduction pathways have already been identified (14–19). The pathways from cell-surface receptors to several prestore and preptsite genes appear to diverge (ref. 20; B.H., J. Pavlovic, and R.P.D., unpublished data). The regions on some genes that are required for regulation by the stimulus of cAMP have recently been identified (21–23).

Reversible protein phosphorylation appears to be involved both in myosin and receptor activation (24–26) and in DNA binding (S.R. Boduluri and R.P.D., unpublished data). Although protein kinase A and C (27, 28) activities have been detected in Dictyostelium, and the gene for the regulatory subunit of protein kinase A (29) has been identified, the genes for the catalytic subunit of Dictyostelium protein kinases have yet to be identified. We have used PCR (30) to synthesize and amplify fragments of catalytic subunits of protein kinase genes. These fragments have allowed us to identify five different members of a protein kinase multigene family and to sequence one full-length cDNA clone, Dd kinase-2. Dd kinase-2 appears to be a single copy gene encoding two transcripts, the larger one being induced during development and by exogenous cAMP.

MATERIALS AND METHODS

DNA Amplification. Dictyostelium genomic DNA was sheared to an average size of 5.0–10 kilobases (kb) for use as the template in the amplification reactions. Oligonucleotide primers were synthesized on an Applied Biosystems oligonucleotide synthesizer. The sequences of the primers were 5′-AGAAGCTTAGT(C/T)GMNNCC(A/G)TGG(A/G)AA-3′ (forward) and 5′-GAGAATTCTGGA(T/G)GCTA(A/G)GTAT(T/C)TGGA(T/G)GTAT(A/G)TC-3′ (reverse). PCR was done with a Coy Tempcycler with a Geneamp kit (United States Biochemical/Cetus) for 35 cycles (1 min at 94°C, 1 min at 45°C, and 3 min at 72°C). Amplified DNA was ligated into the plasmid vector plTZ18U (United States Biochemical) and used to transform Escherichia coli strain JM109. The sizes of inserts in individual clones were determined by digestion with EcoRI and HindIII, electrophoresis on 3% Nuseive/1% Seakam agarose gel, and staining with ethidium bromide.

cDNA Cloning and Nucleic Acid Analysis. A Agt11 cDNA library was provided by Chi-Hung Siu (32) using poly(A) RNA from NC4 cells developing for 8 hr. It was screened with labeled PCR-amplified DNA (33) by hybridization at low stringency. Sixty clones appeared positive and one contained a 2.0-kb cDNA insert that was subsequently amplified by PCR, subcloned into plasmid pTZ18U, and sequenced completely from a series of nested deletions constructed by using Erase-a-Base (Promega) (34, 35). Growth and development of Dictyostelium discoideum AX3 and analysis of DNA and RNA were described (10, 22, 36).
RESULTS

Identification of DNA Fragments Encoding Protein Kinases.
Two regions containing conserved amino acid sequences of serine/threonine kinases (3, 5, 6) were chosen for designing primers for PCR analysis. In designing both primers, the high A+T content and the frequency of codon usage in Dictyostelium (37) were taken into consideration. The sequence of the forward primer derived from the amino acid sequence RDLKPEN was 256-fold degenerate. A HindIII restriction site was attached to the 3' end. The reverse primer derived from the amino acid sequence GTPELYAPE was 288-fold degenerate and was appended to an EcoRI site. PCR was carried out with Dictyostelium genomic DNA and the amplified products were analyzed by electrophoresis on agarose gels (data not shown). A broad band of DNA fragments 140–165 nucleotides long was observed because most typical protein kinases contain 28–35 amino acids between the two conserved regions. No amplified products appeared in the absence of genomic DNA, indicating that the amplification was dependent on the template. The heterogeneity of the PCR product was confirmed by size analysis of individual fragments cloned into a plasmid vector, pTZ18U, and by sequence analysis (34), which identified five different kinase-like sequences, Dd kinase-1 to -5 (Fig. 1). All of these clones encode the invariant tripeptide sequence DFG at approximately the same location as in other typical protein kinases relative to the regions encoded by the forward and reverse primer sequences (Fig. 1). The three hydrophobic amino acids immediately following the conserved asparagine encoded by the forward primer sequence and a cysteine immediately preceding the glycine encoded by the reverse primer sequence are also highly conserved. A threonine located 3 amino acids preceding the glycine encoded by the reverse primer sequence was shown to be phosphorylated in protein kinase A, and this phosphorylation was essential for its interaction with the R subunit (38). Each of the five kinases has a similar phosphorylatable residue at this position: threonine in kinases 1, 2, 3, and 5, but serine in kinase 4. A search through the National Biomedical Research Foundation protein data base with individual Dd kinase sequences revealed homology mainly to other protein kinase sequences, with Dd kinase-2, -3, -4, and -5 peptides showing relatively high homology (50–75%) and Dd kinase-1 showing less homology (40–55%). While Dd kinase-5 showed high homology (≈75%) to protein kinase C, Dd kinase-4 is similar to the protein kinase A (≈75%). Since the short cloned regions encode well-conserved amino acid sequences in protein kinases, it is difficult to deduce their identities by sequence comparison across this region with known protein kinases. The A+T content (≈70%) and the codon usages within the amplified regions are similar to other Dictyostelium genes (data not shown).

---

Molecular Cloning and Analysis of a Protein Kinase cDNA.
Sixty hybridizing clones were identified by screening a Agt11 cDNA library with 32P-labeled PCR products and are being characterized further. Analysis of one of these, Dd kinase-2, by PCR revealed that it contained an ≈2.0-kb insert that hybridized to a 2.2-kb developmentally regulated mRNA. Fig. 2 shows the complete nucleotide sequence and the deduced amino acid sequence for Dd kinase-2. The 1978-base-pair (bp) cDNA sequence contained a single large open reading frame from nucleotides 511–1947 followed by a termination codon TAA and an additional 21 adenine residues. The only other ATG in the 5' untranslated region at nucleotide 154 is followed by several in-frame stop codons within the next 60 bases. Although much longer in comparison to the 5' untranslated sequences of other Dictyostelium mRNAs, the 511 bases of untranslated leader RNA are encoded by A+T-rich DNA typical of Dictyostelium DNA (≈88% A+T) that does not encode protein and contained a CA or CT repeat and an AAT repeat (Fig. 2). The 3' end, a TAA stop codon, contained a 21-nucleotide poly(A) stretch and no polyadenylation signals.

The long open reading frame encodes a protein of 479 amino acids containing two domains: a C-terminal 350-amino acid catalytic domain and a highly basic 130-amino acid N-terminal domain. A search in the National Biomedical Research Foundation protein data base revealed homology only to protein kinases, with highest similarities to protein kinase A, the suppressor of cdce25 of Saccharomyces cerevisiae, S-6 kinase of Xenopus, and protein kinase C (Table 1). The deduced amino acid sequence of Dd kinase-2 contains all the features characteristic of protein kinases and in the same relative positions (Fig. 2) (40): Gly355–Xaa–Gly357–Xaa–Gly359, which forms part of the ATP binding region and also occurs in Dd kinase-2 (G166GKSFG167) [the three-letter amino acid superscripted numbers are derived from bovine cAMP-dependent protein kinase (40) and the single-letter amino acid superscripted numbers are derived from Dd kinase-2]; the Lys72K182, Asp166D276, Asn171N281, all implicated in ATP binding and catalysis; the Asp148D294 and Cys199C311, modifications of which affect catalytic activity; and the tripeptides Asp184 Phe185 Gly186, Gly294DFG296, and Ala200 Pro207, Glu206APE318–320, diagnostic of the protein kinase catalytic domain. In addition, the residue corresponding to Thr193T199 in cAMP-dependent protein kinase is known to be autophosphorylated, and was shown to be essential for interaction with R subunit (40), is also conserved in Dd kinase-2. The N-terminal 130 amino acids are extremely basic (45 basic/11 acidic amino acids) and the DNA encoding this region contained two kinds of repeat sequence motifs previously identified in the coding regions of Dictyostelium genes (39, 41). A CAA repeat in this gene encodes polyglutamine residues and an ACACCA repeat encodes alternating threonine/proline residues. The first 130 amino

---

FIG. 1. Alignment of deduced amino acid sequences of Dictyostelium protein kinases with those of other protein kinases. The deduced amino acid sequences of Dd kinase-1 to -5 were aligned to maximize homology. These were compared with sequences from similar regions of five other well-characterized protein kinases from the National Biomedical Research Foundation data base. The sequences of the primer regions are shown in boldface. Residues identical in 7 or more of the 10 sequences shown are indicated by an asterisk.
Fig. 2. Nucleotide sequence and derived amino acid sequence of Dd kinase-2. The 2.0-kb CDNA was sequenced in both orientations from nested deletion clones. Sequences used as primers for PCR are stippled. The AAAT repeat and the CA/CT repeat in the 5′ untranslated sequence are underlined. The amino acid residues characteristic of protein kinases are in boldface and underlined.

acids are homologous to several proteins containing polyglutamine, including homeobox-containing proteins, and the trans-activator domain of the glucocorticoid receptor.

Regulation of Expression of the Dd Kinase-2 Gene. To determine whether the transcripts from the Dd kinase-2 gene are developmentally regulated, RNA isolated from cells developed for different times was analyzed on Northern blots as described (36). Lanes contain equal amounts of RNA determined by absorbance, intensity of ribosomal RNA bands upon ethidium bromide staining, and, in some experiments, hybridization with probe p74 that detects a gene not regulated by cAMP. Fig. 3 shows hybridization with a 446-bp 3′ end fragment (from nucleotides 1411–1778 of the sequence in Fig. 2) to two differentially regulated mRNAs of 2.0 and 2.2 kb. The 2.0-kb mRNA is expressed maximally in vegetative cells and minimally (37% at 5 hr) in development. The pattern of expression at other times is less than in vegetative cells (40–70%). The 2.2-kb mRNA is not present in vegetative cells. It first appears at the 5-hr stage (30% maximum) and reaches a maximum at the 10-hr stage and thereafter is

Table 1. Homology of Dd kinase-2 to other protein kinases

<table>
<thead>
<tr>
<th></th>
<th>PKA</th>
<th>PKC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sup. of cdc25</td>
<td>39</td>
<td>36</td>
</tr>
<tr>
<td>S-6 kinase</td>
<td>44</td>
<td>55</td>
</tr>
</tbody>
</table>

The sequence of the kinase domain of Dd kinase-2 (amino acids 110–379) was compared to that of bovine protein kinase A (PKA), rat protein kinase C β type (PKC), S. cerevisiae suppressor of cdc25 (Sup. of cdc25), and Xenopus ribosomal protein S-6 kinase (S-6 kinase). The sequences from the National Biomedical Research Foundation protein data base were compared by using the GAP program of University of Wisconsin Genetics Computer Group (49). Percentage identity is indicated in boldface and percentage similarity is in parentheses.

expressed at slightly lower levels (50–70% maximum) throughout development. A probe from the 5′ untranslated sequence (bp 1–500 from the sequence in Fig. 2), a 30- bp antisense oligonucleotide for the sequence from nucleotides 1434–1463, and a full-length cDNA probe all showed identical patterns of hybridization (data not shown). The cDNA contained no EcoRI, Pst I, or EcoRV sites but the sequence predicts single Ava II, Dde I, Taq I, Nco I, HincII, and Hind III sites. Digestion of genomic DNA with several of these enzymes and Southern blot hybridization with a full-length probe showed the number of bands expected with each of these enzymes for a single-copy gene (Fig. 4A). For example, EcoRV (lane R), which does not cut within the gene, gives a single band with all these probes (Fig. 4 A–C), whereas Dde I (lane D) gives two bands with the full probe (Fig. 4A) and one each with the 5′ and 3′ probes (Fig. 4 B and C). Analysis with either the 5′ end (Fig. 4B) or the 3′ end (Fig. 4C) probe shows unique bands hybridizing to these probes and allows us to construct a partial restriction map of the genomic DNA (Fig. 4D). In addition, under low-stringency conditions, the full-length probe shows weak hybridization to several fragments, presumably repeated sequences and/or other kinase genes (data not shown).

Several postaggregation-stage genes were shown to be regulated by exogenous cAMP (36, 42, 43) through cell-surface cAMP receptors (9–13). These experiments were

![Fig. 3. Developmental regulation of Dd kinase-2. RNA samples isolated from AX3 cells developed on filters and harvested at the indicated times were hybridized with labeled 3′ end fragment (see text) and the cDNA insert. Time 0 represents RNA from vegetative cells. The two RNA species (∼2.0 and 2.2 kb) are indicated by arrows.](image-url)
DISCUSSION

We have identified five different members of a protein kinase multigene family of Dictyostelium from analyzing only 10 cloned DNA fragments that were amplified by PCR using primers designed from conserved sequences. Thus, we believe that other different sequences may be present in the collection of our cloned PCR fragments. Because of our choice of primers, we expect to amplify only protein-serine/threonine kinases, and indeed all the clones amplified show more homology to protein-serine/threonine kinases than to protein-tyrosine kinases. They also contain a phosphorylatable threonine or serine residue at approximately the same position as the autophosphorylatable threonine in other protein-serine/threonine kinases.

While we have not shown that Dd kinase-2 protein has protein kinase activity, the high degree of homology with protein kinases and the absolute conservation of amino acid residues known to be involved in kinase activity strongly suggest that this is indeed a protein kinase gene, a criterion now commonly accepted (3–7). Based on homology, Dd kinase-2 may be classified under the subgroup of cyclic nucleotide and calcium phospholipid-dependent protein kinases described by Hanks et al. (3). However, because of the basic sequence in the N-terminal region of the protein, it may represent a unique member of that subgroup. By analogy with other protein kinases, the N-terminal domain may play a role in the regulation of kinase activity and/or localization of the protein to specific intracellular locations.

Of the two mRNAs detected, the 2.2-kb mRNA appears only after cells establish a cAMP signaling system and is expressed when tight aggregates are formed. Since exogenous cAMP but not bromo-cAMP increases its expression, this gene is likely regulated by cell-surface cAMP receptors. Because this mRNA accumulates relatively early compared...
Developmental Biology: Haribabu and Dottin

with several other cell-type-specific mRNAs regulated by cAMP, we surmise that Dd kinase-2 may play some role in later events in signal transduction. The pattern of expression of the two mRNAs suggests that the kinase may have specific functions in both growing and developing cells. The likely presence of only one copy of the gene and two mRNA species suggests that differentially regulated multiple transcripts must originate from separate promoters and/or from alternative splicing events as was observed in the cyclic nucleotide phosphodiesterase gene of Dictyostellium (44, 45). The unusually long 5’ untranslated sequence (>500 nucleotides) of Dd kinase-2 mRNA suggests the possibility of further regulation at the translational level (45–47).

In summary, our results suggest the existence of a large protein kinase multigene family in Dictyostellium, a system amenable to molecular genetic analysis. Using a similar approach involving PCR, we have also identified fragments of type I and type IIA phosphoprotein phosphatase DNAs of Dictyostellium (48). Thus, it may be possible not to identify the specific protein kinases required for growth and development of eukaryotes, but also to define specific pathways within which these protein kinases and protein phosphatases function to mediate differentiation and/or development.

We thank Dr. Chi-Hung Siu for generously providing the Agt11 bank, Mr. Steve Bobin for timely synthesis of oligonucleotides, Ms. Rebecca M. Crespo for expert typing and preparation of figures, Mrs. Geraldine Reichel for technical assistance, and Dr. Joseph Krakow for critical reading of the manuscript. This work was supported by National Institutes of Health Grant GM27310-11 to R.P.D. and an award (RR-03037) to Hunter College, City University of New York, for the support of a Center for Study of Gene Structure and Function by the National Institutes of Health–Research Centers in Minority Institutions Program.