The MsK family of alfalfa protein kinase genes encodes homologues of shaggy/glycogen synthase kinase-3 and shows differential expression patterns in plant organs and development

Aniko Pay1,², Claudia Jonak1, Laszlo Bőgre1,², Irute Meskiene1,², Theresia Mairinger1, Aladar Szalay2, Erwin Heberle-Bors1 and Herbert Hirt1,a

1Institute of Microbiology and Genetics, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, Austria,
2Plant Molecular Genetics, 6–30 Medical Science Building, University of Alberta, Edmonton, Canada T6G 2H7, and
3Institute of Botany, Lithuanian Academy of Science, Vilnius 2039, Lithuania

Summary

This paper reports on the isolation of a novel class of plant serine/threonine protein kinase genes, MsK-1, MsK-2 and MsK-3. They belong to the superfamily of cdc2-like genes, but show highest identity to the Drosophila shaggy and rat GSK-3 proteins (65–70%). All of these kinases share a highly conserved catalytic protein kinase domain. Different amino-terminal extensions distinguish the different proteins. The different plant kinases do not originate from differential processing of the same gene as is found for shaggy, but are encoded by different members of a gene family. Similarly to the shaggy kinases, the plant kinases show different organ-specific and stage-specific developmental expression patterns. Since the shaggy kinases play an important role in intercellular communication in Drosophila development, the MsK kinases are expected to perform a similar function in plants.

Introduction

In Drosophila, the segment-polarity genes are required for the establishment of the metameric pattern of the embryo, specifying positions of cells within the segments (Nüsslein-Volhard and Wieschaus, 1980). The number, position and polarity of the metameric units is dictated by the products of interacting maternal and zygotic segmentation genes. The activity of the segment-polarity proteins has been suggested to control the cellular interactions which convey the positional information to the involved cells (Martínez-Arias et al., 1988). The predicted protein sequences of some of the segment-polarity genes are consistent with a function in a signal transduction process involved in the determination of cell identity within the segmental unit (Hooper and Scott, 1989; Nakano et al., 1989; Van den Heuvel et al., 1989).

Intercellular interactions also appear to be involved in the determination of cell fates during Drosophila development. For example, the central nervous system arises from a number of neuroblasts that segregate from initially ectodermal cells. If a neuroblast is removed experimentally, a neighbouring cell will change its fate and become a neuroblast (Doe and Goodman, 1985). A process called lateral inhibition normally prevents neighbour cells from becoming neuroblasts. Lateral inhibition is thought to be mediated by a signal transduction pathway that converts the extracellular signals to changes in gene expression. The protein sequences of several neurogenic genes involved in these processes are compatible with such a function (Campos-Ortega, 1988).

Drosophila shaggy mutants show two seemingly unrelated phenotypes with characteristics similar to both of the above described processes. One type of shaggy mutants results in the segregation of excess numbers of neuroblasts in the embryo and bristle mother cells in the imago (Bourois et al., 1989; Simpson et al., 1988). Functional shaggy has been shown to be necessary for the intercellular communication process which mediates the formation of only one bristle mother cell (Simpson and Carteret, 1989). The other type of shaggy mutants form abnormal embryos in that all cells develop neural characteristics (Bourois et al., 1989). The molecular identification of shaggy has revealed that the gene encodes several distinct serine/threonine protein kinases which are developmentally regulated (Bourois et al., 1990; Siegfried et al., 1990). Although the different kinases have identical catalytic domains, different amino-terminal extensions are added by differential processing.

Shaggy shows 75% identity to glycogen synthase kinase-3/Factor A (Woodgett, 1990). Glycogen synthase kinase-3 (GSK-3) was first identified as a protein kinase that phosphorylates glycogen synthase (Embi et al., 1980; Hemmings et al., 1982). GSK-3 was also shown to be identical to Factor A, the activator protein of protein phosphatase-1 (Hemmings et al., 1982; Vandenhedee et
The open reading frames of all MsK clones were very similar. Overall, identities of 83% for MsK-1/2, 88% for MsK-2/3 and 83% for MsK-1/3 were obtained. Although the majority of amino acid differences were found at the N- and C-terminal ends, the presence of conserved domains within these regions was still observed. Analysis of the predicted amino acid sequences revealed that all three genes contain the conserved sequence domains of serine/threonine protein kinases (Hanks et al., 1988).

MsK-1, MsK-2 and MsK-3 are homologues of GSK-3

A homology search in current protein sequence databanks showed highest similarity to the rat GSK-3 and to the Drosophila shaggy proteins. All predicted MsK proteins were found to have identities of 65-70% to different members of the GSK-3 and shaggy families. Significant homology was also observed to the cdc2 kinases (35-40%) from yeast (Hindley and Phear, 1984; Lörrincz and Reed, 1984), humans (Lee and Nurse, 1986), and plants (e.g. Hirt et al., 1991) and to the yeast protein kinases KSS1 (Cournede et al., 1989), Fus3 (Ellson et al., 1990) and Kin28 (Simont et al., 1986) (33, 32 and 34%, respectively) and to extracellular signal-regulated kinase-1 (ERK-1; Boulton et al., 1990) with 31% identity. All of these kinases belong to the superfamily of cdc2 genes. Comparison of the predicted MsK, GSK-3 and shaggy proteins with the cdc2 kinases revealed that all kinases have a similar arrangement of the nucleotide binding region and the catalytic kinase core domain. However, the PSTAIRE motif, which is conserved in all cdc2 kinases contains interspersed between these two domains, is missing in GSK-3, shaggy and MsK-type kinases. Moreover, the GSK-3, shaggy and MsK kinases contain N-terminal extensions not found in the cdc2 kinases.

Shaggy and GSK-3 proteins are highly similar (identities between 73 and 78% for different pairs of the family members). The divergence between the two members of the GSK-3 family has been found to arise from the presence of different amino-terminal sequences (Woodgett, 1990). In Drosophila, seven different shaggy transcripts are produced, each at a specific stage in embryogenesis (Bouros et al., 1990; Siegfried et al., 1990). Comparison of maternal and zygotic shaggy cDNA clones suggests that the N-terminal divergence arose through differential processing of 5' exons (Siegfried et al., 1990). Alignment of the predicted MsK-2 amino acid sequence with maternal shaggy and GSK-3β reveals that the sequences are very similar over nearly the entire length, excluding the amino-
Figure 1. Nucleotide and predicted amino acid sequences of the alfalfa MsK CDNA clones.
(a) MsK-1; (b) MsK-2; (c) MsK-3.
Alfalfa protein kinases homologous to shaggy/GSK-3

Figure 1. Continued.
terminal regions as shown in Figure 3. Also within the Msks, homology is much reduced in the amino-terminal regions (Figure 2). Including these divergent amino-terminal regions, amino acid identities between the Msks and shaggy/GSK-3 kinases amount to scores between 65 and 70% for different pairs.

**Highly different transcript levels of MsK-1, -2 and -3 in vegetative alfalfa organs and suspension culture cells**

To determine the size of the mRNAs of the MsK-1, -2 and -3 genes, Northern analysis was performed with poly(A)+ RNA from alfalfa suspension cultures. To avoid crosshybridization, 3’ end specific segments were synthesized by PCR. MsK-1, -2 and -3 genes were found to encode mRNAs of approximately 1.6 kb, 1.5 kb and 1.8 kb, respectively. Whereby MsK-1 was found to be strongly expressed, MsK-2 and MsK-3 transcript levels were very low (data not shown).

To analyse the expression of the three MsK genes in various vegetative organs of alfalfa plants, RNA was prepared from leaves, petioles, stems, roots and nodes. Northern analysis with both total RNA or poly(A)+ RNA resulted in detection of constitutive transcript levels of MsK-1. However, no transcripts could be observed for MsK-2 or MsK-3, even after 2 weeks of exposure (data not shown). Since we knew from Northern analysis of suspension cultured cells that MsK-2 and MsK-3 transcript levels were very low, transcript levels of all three genes were quantitatively determined by reverse transcribing the RNA samples and PCR amplification with specific primers. The following controls were performed. First, to certify that equal amounts of total RNA were used in the reverse transcription reactions, the RNA samples were quantified by spectrophotometry and by ethidium bromide staining of denaturing agarose gels, as shown in Figure 4a. Second, all reactions were performed at the same time in parallel. Third, to test whether the RNA preparations contained contaminating genomic DNA traces, PCR was performed with and without the use of the
The distribution of MsK transcripts isolated from alfalfa leaves (lanes 1 to 5, respectively). The transcripts were amplified by reverse transcription, separated on agarose gels, blotted to nylon filters, and hybridized with gene-specific probes. The result of such an analysis is shown in Figure 5. MsK-1 was found to be constitutively expressed in all five organs (Figure 5a, lanes 1 to 5). In contrast, MsK-2 was absent in leaves and petioles but expressed in stems, roots, and nodes (Figure 5b, lanes 1 to 5, respectively). MsK-3 transcripts were completely absent in leaves and petioles, hardly detectable in stems, and moderately expressed in nodes (Figure 5c, lanes 1 to 5, respectively). The amplification products of the control cdc2 gene (Figure 5d) show that the differences in the MsK transcript levels are reflected in the steady-state levels of the mRNAs in different organs.

Pattern of MsK-1, -2, and -3 transcript levels during flower and seed development

To analyze whether MsK-1, -2, and -3 genes may play a specific role during flower and seed development, several organs were quantitatively analysed. For this purpose, total RNA was isolated from flower buds (corolla plus the sepals), developing flowers (corolla prom from sepals), unopened flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-1 levels were only detected in the developing flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-2 levels were only detected in the developing flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-3 levels were absent in all organs (lanes 1 to 5 in Figure 6, respectively). The qPCR results for the control cdc2 gene (Figure 6d) show that the differences in the MsK transcript levels are reflected in the steady-state levels of the mRNAs in different organs.

Figure 4: Reverse transcribed PCR amplification.
(a) To prove that equal amounts of RNA were used per organ, total RNA was isolated from alfalfa leaves, stems, roots, nodes, flower buds, developing flowers, unopened mature flowers, green seed pods and globular, heart-shaped and torpedo-shaped somatic embryos (lanes 1 to 11, respectively). After spectrophotometric quantification, 10 µg of each sample was loaded onto a 1% denaturing agarose gel and stained with ethidium bromide. (b) To prove that reverse transcription is necessary for PCR amplification of transcripts, total RNA was isolated from suspension cultured alfalfa cells and treated with or without DNase (lanes 1, 2, 3, and 4, respectively) and with or without reverse transcriptase before amplification with PCR (lanes 1, 3, and 2, 4, respectively). (c) To prove linearity of the amplification system, 0.1, 0.5, and 1 µg of total RNA was used for each reaction (lanes 1 to 3, respectively). PCR products for the cdc2 gene were separated on a 2% agarose gel, blotted to nylon filters and hybridized with a radiolabeled cdc2 fragment.

Figure 5: Organ distribution of MsK transcripts.
Total RNA was isolated from alfalfa leaves (lanes 1 to 5), stems, roots, nodes, flower buds, developing flowers, unopened mature flowers, green seed pods and globular, heart-shaped and torpedo-shaped somatic embryos (lanes 1 to 11, respectively). After spectrophotometric quantification, 10 µg of each sample was loaded onto a 1% denaturing agarose gel and stained with ethidium bromide. MsK-1 (a) MsK-2 (b) MsK-3 (c) cdc2 (d)

For this purpose, total RNA was isolated from alfalfa leaves (lanes 1 to 5), stems, roots, nodes, flower buds, developing flowers, unopened mature flowers, green seed pods and globular, heart-shaped and torpedo-shaped somatic embryos (lanes 1 to 11, respectively). After spectrophotometric quantification, 10 µg of each sample was loaded onto a 1% denaturing agarose gel and stained with ethidium bromide. MsK-1 (a) MsK-2 (b) MsK-3 (c) cdc2 (d)

Changing pattern of MsK-1, -2, and -3 during flower and seed development

To analyse whether MsK-1, -2, and -3 genes may play a specific role during flower and seed development, several organs were quantitatively analysed. For this purpose, total RNA was isolated from flower buds (corolla plus the sepals), developing flowers (corolla prom from sepals), unopened flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-1 levels were only detected in the developing flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-2 levels were only detected in the developing flowers and green seed pods (lanes 1 to 5 in Figure 6, respectively). MsK-3 levels were absent in all organs (lanes 1 to 5 in Figure 6, respectively). The qPCR results for the control cdc2 gene (Figure 6d) show that the differences in the MsK transcript levels are reflected in the steady-state levels of the mRNAs in different organs.
Figure 6. MsK-1, -2 and -3 transcript levels during alfalfa flower and seed development.
Total RNA was isolated from alfalfa flower buds (lane 1), immature developing flowers (lane 2), unopened mature flowers (lane 3), fully opened flowers (lane 4) and green seed pods (lane 5). After first strand synthesis, MsK-1, -2, -3 or cdc2 transcripts were amplified with specific primers by PCR. PCR products were separated on 2% agarose gels, blotted to nylon membranes and hybridized to 3’ probes specific for: (a) MsK-1; (b) MsK-2; (c) MsK-3; (d) cdc2.

Figure 7. MsK-1, -2 and -3 transcript levels in somatic embryos.
Total RNA was isolated from three fractions of somatic embryos: globular-shaped embryos (lane 1), heart-shaped embryos (lane 2) and torpedo-shaped embryos (lane 3). After first strand synthesis, MsK-1, -2, -3 or cdc2 transcripts were amplified with specific primers by PCR. PCR products were separated on 2% agarose gels, blotted to nylon membranes and hybridized to 3’ probes specific for: (a) MsK-1; (b) MsK-2; (c) MsK-3; (d) cdc2.

MsK-1, -2 and -3 transcripts in somatic embryos

In Drosophila, shaggy is developmentally transcribed during embryogenesis, giving rise to different mRNAs in each stage. To analyse the situation of the MsK genes in plants, we determined mRNA levels of the three MsK genes in various stages of alfalfa somatic embryos: globular embryos (lane 1), heart-shaped embryos (lane 2) and torpedo-shaped embryos (lane 3). MsK-1 transcript levels did not change appreciably (Figure 7a). MsK-2 transcript levels could only be observed in globular embryos (Figure 7b, lane 1) while MsK-3 transcripts were moderate in globular embryos and decreased in the later stages (Figure 7c, lanes 1 to 3, respectively).

Discussion

Shaggy belongs to the class of segment-polarity genes, necessary for correct pattern formation within each segment. It also plays a role in epidermal cell differentiation in the adult fly (Bourouis et al., 1989; Simpson et al., 1988). Mutations in segment-polarity genes lead to deletions of duplications of parts in each segment. They are thought to act co-ordinately by means of intercellular communication. The transcription factors c-jun and c-myb, glycogen...
shown). Finally, all three MsK genes were found to give rise to single transcripts in Northern analyses.

Comparison of the predicted protein sequences of the two mammalian GSK-3 genes and the two characterized Drosophila shaggy transcripts revealed high conservation over most of the sequence. However, sequence similarity of the two GSK-3 proteins disappeared in the N-terminal direction at the same conserved amino acid motive where the differentially processed shaggy gene products diverge from each other. Alignment of the MsK proteins with the GSK-3 and shaggy sequences revealed a similar picture. The N-terminal protein sequences of all three MsK genes were very different and showed no similarity to either the shaggy or the GSK-3 proteins. Homology started from the same conserved amino acid domain as was found in the comparison of the Drosophila and mammalian kinases.

All serine/threonine protein kinases including the cdc2-like kinases contain eleven conserved domains (Hanks et al., 1988) which are found in the carboxyl part of all shaggy, GSK-3 and MsK protein kinases. These proteins thereby resemble amino-terminally extended versions of a catalytic kinase subunit. These extensions appear to be conserved during evolution since GSK-3p and the ovarian product of shaggy have highly similar amino termini and must be considered to be true functional homologues of each other.

Since the catalytic domain carries all the relevant sequence domains for a functional protein kinase, the N-terminal modules may have a regulatory function. Each amino-terminal module may interact with specific proteins that couple the catalytic core module to different regulators. Alternatively, different N-termini could induce different conformational states of the respective kinases resulting in an altered substrate specificity. Taken together, attachment of different amino terminal sequences to a common serine/threonine protein kinase domain might mediate external signals to a variety of different substrates, including activation of transcription factors that switch on particular target genes involved in a particular developmental or metabolic pathway.

In Drosophila, different shaggy transcripts were found to be expressed in a developmental stage-specific manner. Similarly, the alfalfa MsK genes were found to be transcribed in an organ- and developmental stage-specific pattern. Therefore, it is an intriguing possibility that the MsK genes could perform analogous functions in plants to those which the shaggy transcripts do in insects. In particular, the highly regulated pattern of the MsK-2 and MsK-3 genes during somatic embryogenesis is reminiscent of the short-lived embryonic forms of shaggy. In contrast, the constitutive expression of the MsK-1 gene is suggestive of a late developmental role similar to the shaggy form which is responsible for the epidermal differentiation in the adult fly.

To investigate the role of the shaggy-like kinases in plants, the study of the substrate specificity of each MsK kinase will be indispensable. For this purpose, we are presently producing peptide antibodies against the amino-terminus of each kinase. Furthermore, although we have obtained clear evidence that the MsK genes are differentially transcribed in various vegetative organs and during flowering, it will be important to identify the particular tissues or cells that actively synthesize the kinases. Finally, by expression of mutant versions of the MsK genes, we hope to interfere with their normal function in plants. It will be seen whether, in analogy to Drosophila or mammals, embryogenesis, organogenesis or specific metabolic pathways will be affected by this approach.

**Experimental procedures**

**Plant tissue culture**

A suspension culture of Medicago sativa L. (genotype RA3) was used (György et al., 1991). The suspension culture was derived from callus tissue which originated from stem segments and was grown in Schenk and Hildebrandt (1972) medium supplemented with 15 μM naphthalene acetic acid and 10 μM kinetin. Induction of somatic embryos was performed as described (György et al., 1991).

**Library screening and sequence analysis**

Poly(A)+ RNA was isolated from somatic embryos (György et al., 1991). A lambdaZAP-cDNA library was prepared according to the manufacturer (Stratagene, La Jolla, CA). Two degenerate oligonucleotides were synthesized: a (5'GAA TTC GGN GAA/G GGI ACI TAT/C GGN GT-3') and b (5'-AAG CTG TGG TCG TIA/G AVTCA NCN/TA/GTG-3'). They correspond to two highly conserved regions of the superfamily of cdc2 kinases and encode either the nucleotide binding domain (gly-glu-glu-thr-tyr-gly-val) or domain V of the catalytic domain (his-arg-asp-ieu-lys-pro-gin-ASN-leu). After gel elution from 10% denaturing polyacrylamide gels, 50 ng of each batch of oligonucleotides was 32P-labelled with 50 μCi gamma-ATP and polynucleotide kinase. Approximately 300 000 colonies were screened with both probes. Putatively positive clones were rescreened to a minimum fit of 16 nucleotides according to Wood et al. (1985). The only positive clone was sequenced from both directions with the T7 polymerase sequencing kit (Pharmacia, Uppsala). The insert from this clone was used as a probe to rescreen the library. Fifteen positive clones were identified. They could be classified into three groups, called MsK-1, -2 and -3. The longest members of each group were sequenced as above.

**RNA extraction and PCR quantification of transcript levels**

RNA from suspension culture cells or plant tissue was prepared as described (Hirt et al., 1991). The quantitation of MsK-1, -2 and -3 mRNA levels was principally done according to Scherer et al. (1990) with the modification that a DNase treatment of the total RNAs was included. This found to be necessary since PCR amplification was obtained also when no reverse transcriptase was used. However, after DNase treatment, no amplification was obtained. For this purpose, 5 μg total RNA were digested for 30 min at 37°C with 3 units of RNase-free DNase (Pharmacia) in 40 mM Tris-HCl, pH 7.5, 6 mM MgCl2, containing 5 units of RNAsin (Sigma). After incubation, the DNase was inactivated by
heating the reaction mixture for 3 min at 65°C. The 3' primers for healing the reaction mixture for 3 min at 65°C. The 3' primers for the reverse transcription were: for 3' MsK-1 (CAT CTA GAA CGT- TCA TAT GGA A), for 3' MsK-2 (TAG AAT TCA CAA TGC TCC TAG T), for 3' MsK-3 (AAT TCG AGT GAA CAC ATG AGA T) and for 3' cdc2 (GCC TGC AGC AAA CTG AAG TCC CT). Annealing of the primers to 0.5 μg of DNase-treated RNA was done by heating at 95°C for 2 min, by annealing at 60°C for 1 min and by cooling to 37°C. PCR was performed with the above 3' end primers and the following 5' end primers: for 5' MsK-1 (AGG AAT TCC AGT AGC GAT TCT G), for 5' MsK-2 (ATA AGC TTA GAG TGC TTT G), for 5' MsK-3 (TGG TGC ACG CTT GTG TAG TTA TTA) and for 5' cdc2 (GGA AGC TTA ATG GTA AAG GTG T). PCR was done for 20 cycles of 1 min at 92°C, 2 min at 61°C and 3 min at 72°C. The amplified fragments were run on 2% agarose gels, blotted to nylon membranes and hybridized to 35S-labelled 3'-specific MsK-1, MsK-2, MsK-3 and cdc2 probes.

Acknowledgements
We thank Zarko Harzenjak for technical assistance. This work was supported by grants S6004 BIO and P8883-MCB from the Österreicherischen Fonds zur Förderung der Wissenschaften.

References


EMBL Data Library accession numbers X68411 (MsK-1), X68410 (MsK-2) and X68409 (MsK-3).