A novel alpha-helical protein, specific to and highly conserved in plants, is associated with the nuclear matrix fraction

Annkatrin Rose³, Frank Gindullis¹,³ and Iris Meier²

Plant Biotechnology Center and Department of Plant Biology, Ohio State University, Columbus, OH 43210, USA

Received 19 August 2002; Accepted 30 November 2002

Abstract

A cDNA for a novel plant protein was isolated from tomato. Nuclear Matrix Protein 1 (NMP1) is a ubiquitously expressed 36 kDa protein, which has no homologues in animals and fungi, but is highly conserved among flowering and non-flowering plants, including gymnosperms, moss, and the liverwort Marchantia polymorpha. NMP1 is predominantly α-helical with multiple stretches of short amphipathic regions. Cell fractionation, immunofluorescence, and GFP localization experiments showed that NMP1 is located both in the cytoplasm and nucleus and that the nuclear fraction is associated with the nuclear matrix. NMP1 is a candidate for a plant-specific structural protein with a function both in the nucleus and cytoplasm.

Key words: Amphipathic alpha helix, GFP, nuclear matrix, tomato.

Introduction

The nuclear matrix is operationally defined as the insoluble nuclear substructure that remains after the majority of DNA and soluble and chromatin-bound proteins have been removed from the nucleus (Berezney and Coffey, 1974; Nickerson et al., 1997; Wan et al., 1999). Electron micrographs show that the animal nuclear matrix consists of the nuclear pore complexes embedded in the nuclear lamina, and a network of internal 10 nm filaments, into which granular structures and the nucleoli are embedded (Verheijen et al., 1988; Penman, 1995). In two-dimensional protein gels of nuclear matrix preparations, more than 200 polypeptides can be distinguished, but only a few components of the nuclear matrix have been cloned (Fey and Penman, 1988; Stuurman et al., 1990).

Best studied from animal systems are the nuclear lamins, a group of intermediate filament (IF) proteins that form the lamina, a filamentous protein meshwork, lining the nuclear envelope. The nuclear lamins are attached to the membrane by farnesylation and interactions with integral membrane proteins (McKeon et al., 1986; Furukawa et al., 1995). A second, filament-like protein present in the nuclear matrix is NuMA, a protein also localized at the spindle poles during mitosis (Lydersen and Pettijohn, 1980). NuMA has a filament-like secondary structure and induces the formation of a three-dimensional lattice inside the interphase nucleus when overexpressed in Hela cells, indicating that it may play a structural role in the architecture of the nuclear matrix (Harborth and Osborn, 1999). The available data suggest that the nuclear matrix consists of a filamentous nuclear skeleton that is involved in chromatin organization and in different aspects of nucleic acid metabolism. However, with the exception of the nuclear lamins and maybe NuMA, the structure-forming proteins of the nuclear matrix have not yet been identified from any organism.

Plant cells contain a nuclear matrix with similar characteristics to the one in animal cells. It shows a filamentous structure in EM micrographs (Moreno Díaz de la Espina, 1995). Several attempts have been made to identify plant lamins or intermediate filament-like proteins. Plant proteins cross-reacting with antibodies raised against animal lamins and IF proteins have been detected (Frederick et al., 1992; McNulty and Saunders, 1992; Mingué and Morena Díaz de la Espina, 1993). Those plant
IF-like proteins have molecular weights ranging from 58 to 68 kDa, and they have been found not only at the rim of plant nuclei and nuclear matrix fractions, but also distributed throughout the internal nuclear matrix (Frederick et al., 1992; McNulty and Saunders, 1992). A 200 kDa protein that cross-reacts with an anti-NuMA antibody has also been detected in the onion nuclear matrix (Yu and Morena Diaz de la Espina, 1999). Despite their earlier description, no lamin-like proteins or other IF-like proteins have been cloned from plants. Neither the yeast genome nor the Arabidopsis genome sequences (Mewes et al., 2002) appear to contain lamin genes. Therefore, it is conceivable that in non-animal organisms other proteins functionally replace the nuclear lamins. It is possible that the animal anti-IF antisera detect different plant proteins with similar antigenic determinants.

The cloning of a novel plant protein (Nuclear Matrix Protein 1, NMP1) which has no homologues in animals and fungi, but is highly conserved among flowering and non-flowering plants is described here. NMP1 consists mainly of amphipathic alpha-helices, a signature motif of filament-like proteins. It is present both in the cytoplasm and the nucleus, and the nuclear-localized fraction is associated with the operationally defined nuclear matrix.

**Materials and methods**

**Plant material**

Tomato (Lycopersicon esculentum cv. VFNT cherry) and tobacco (Nicotiana tabacum) plants were grown under greenhouse conditions. Tobacco suspension culture cells (BY-2 cells and NT-1 cells) were cultured as described previously (Nagata et al., 1992; Allen et al., 1996). For all experiments aliquots from a mid-log phase culture (4–6 d) were taken.

**Yeast two-hybrid screen**

Yeast two-hybrid screening was performed as described earlier (Gindullis et al., 1999).

**Plasmid vectors**

The plasmid pAD 6-6 contains the complete NMP1 open reading frame and 76 bp 5′-untranslated region. The NMP1 cDNA was cloned as an EcoRI–Khol fragment from pAD 6-6 into pBluescript SK II+ (Stratagene, La Jolla, USA) to create pBS 6-6 EX. To express the NMP1 protein in E. coli, the cDNA insert from pBS 6-6 EX was cloned as a BamHI–KpnI fragment into the BamHI and KpnI sites of pRSETC (Invitrogen) to create pRSETC-NMP1.

To create pNMP1-GFP for in vivo localization experiments, the open reading frame of NMP1 was amplified by PCR using the primers 5′-TGG ACA TGT CAG CGA AAC-′ and 5′-CGA CAT GTT ATC TTC ACC-′ and pAD 6-6 as a template. The resulting PCR product was cloned into the single NcoI site of pRTL2-pmGFPS65T (von Armin et al., 1998) using the internal AflIII sites of the PCR primers.

**DNA hybridization screen**

A young tomato fruit (3–8 mm) λZAP cDNA library (Meier et al., 1996) was screened using the insert of pAD 6-6 in a DNA hybridization screen, essentially as described by Sambrook et al. (1989).

**Protein expression, purification, and antibody production**

Protein expression, purification, and antibody production were done essentially as described by Gindullis et al. (1999).

**Secondary structure analysis, database searches, and sequence alignments**

Secondary structure and charge profile analysis of NMP1 were performed using the protein analysis program and the CLUSTAL algorithm with the default parameters of gap penalty of 10 and gap length penalty of 10.

**Isolation of BY-2 total protein, NT-1 and BY-2 protoplasts, nuclei, and nuclear matrices**

Protoplasts, nuclei, and nuclear matrices from NT-1 and BY-2 suspension-cultured cells were isolated as described by Hall and Spiker (1994). Rat nuclear matrix isolated as described by Mortillaro and Berezney (1998) was kindly provided by Dr Ronald Berezney. For the isolation of BY-2 total protein 1 ml of the protoplast suspension used for the isolation of nuclei was sedimented for 5 min at 300 g. After adding 0.5 ml extraction buffer (62.5 mM TRIS-HCl pH 6.8, 20% glycerol, 4% SDS, 1.4 M β-mercaptoethanol) to the cells, the sample was vortexed for 30–60 s and incubated for 10 min at 70 °C. Debris was removed by centrifugation for 10 min at 4 °C and 15 000 rpm in a tabletop centrifuge. The cleared supernatant was transferred to a fresh tube and stored at −80 °C. The quality of the nuclei and nuclear matrix samples was analysed by microscopy and SDS-PAGE. For SDS-PAGE, nuclei and nuclear matrix samples were boiled for 3 min in 15 μl of SDS-loading buffer (Sambrook et al., 1989) immediately prior to loading and approximately 5–10 μg of total protein were loaded.

**Isolation of total plant protein**

Total plant protein was isolated as described by Harder et al. (2000).

**Immunoblot analysis**

A 1:3 000 dilution of the anti-NMP1 antiserum and a 1:1 000 dilution of horseradish peroxidase-coupled donkey anti-rabbit secondary antibody (Amersham Pharmacia Biotech, Piscataway, NJ) were used to perform immunoblot analyses as described by Sambrook et al. (1989). Enhanced chemiluminescence detection was performed as described by the manufacturer (Amersham Pharmacia Biotech). Protein gels were stained with BluPrint Fast-PAGE Coomassie stain (Life Technologies, Rockville, USA).

**Transformation of BY-2 cells**

Plasmid DNA for transient transformation of BY-2 cells and protoplasts was purified using the Wizard Plus Maxi Prep DNA Purification System (Promega, Madison, WI) according to the manufacturer’s protocol. Ballistic transformation of BY-2 cells was performed as described by Gindullis and Meier (1999).

**Fixation, immunolabelling, and fluorescence microscopy of NT-1 cells and protoplasts**

Fixation, immunolabelling, and fluorescence microscopy of fixed and living cells and protoplasts were performed as described by Gindullis et al. (1999).
RNA gel blot analysis

Total RNA from tomato leaves (5–15 mm in length), fruits (3–8 mm in diameter), whole flowers, and 7-d-old etiolated and light-grown seedlings was isolated with the Trizol Reagent from Life Technologies (Gaithersburg, MD). Total RNA (10 \( \mu \)g each) was separated on a formaldehyde gel, blotted to a nitrocellulose membrane, and hybridized with the radioactively labelled 1031 bp EcoRI fragment of the NMP1 cDNA from pAD12-4 essentially as described by Sambrook et al. (1989).

Results

The cDNA for NMP1 (Nuclear Matrix Protein 1) was originally cloned in a yeast two-hybrid experiment, but the interaction with the bait construct could not be confirmed. The yeast two-hybrid screen was performed with a cDNA expression library made from young tomato leaves (Gindullis et al., 1999). A prey plasmid (pAD 6-6) was identified that led to activation of the \( \text{HIS3} \) and \( \text{lacZ} \) reporter genes in the reporter strain YRG-2. Sequencing showed that the longest open reading frame on pAD 6-6 was not in the reading frame +2 (in frame with the GAL4 activation domain (AD)) but in the reading frame +1. Reinsertion of the correct open reading frame did not lead to activation of the reporter genes. Therefore, the observed yeast two-hybrid interaction was interpreted as 'false positive'. Because of its novel sequence, high degree of conservation among flowering and non-flowering plants, and appearance only within the plant kingdom, NMP1 was
further investigated. The cDNA contained in the yeast two-hybrid prey plasmid pAD 6-6 had a size of 1295 bp. The open reading frame continued to the 5’ end of the cDNA, but an ATG at position 70 showed a sequence environment...
that is similar to the consensus sequence for plant initiation codons \((\text{A ACA ATG GC})\) (Luetke et al., 1987).

To identify longer cDNAs and to test if another ATG is present further upstream, a cDNA library hybridization screen was performed with the insert of pAD 6-6 as the hybridization probe. The longest cDNA isolated was only 27 bp longer than the original pAD 6-6 cDNA (Fig. 1). No additional start codon was present on this sequence. As the size of the longest NMP1 cDNA (1324 bp) was close to the predicted size of the NMP1 mRNA (c. 1350 nucleotides, see Fig. 5), it most likely contains the complete coding capacity for NMP1. This was also confirmed by the alignment between the tomato NMP1 cDNA and the cDNA of the corn and rye NMP1 orthologues. The high degree of similarity between the three cDNA sequences ends immediately upstream of the predicted start ATG (data not shown). The NMP1 open reading frame codes for a putative protein of 331 amino acids (aa) and a calculated molecular mass of 37.4 kDa. Figure 1 shows the predicted amino acid sequence of NMP1. Database searches indicated that NMP1 is a novel protein. It is highly \(\alpha\)-helical and has multiple stretches of amphipathic \(\alpha\)-helices (Figs 1, 2A). It contains six potential phosphorylation sites for protein kinase CKII (underlined in Fig. 1). The charge profile shows alternating basic and acidic domains (Fig. 2A).

Few similarities were found with animal or yeast proteins in a GenBank search and the best matches, although still weak, are the yeast nuclear protein MAK16 (Wickner, 1988) and a cephalopod intermediate filament protein (IF) from Sloane’s squid (Tomarev et al., 1993). Both proteins are, like NMP1, highly \(\alpha\)-helical. MAK16 is the protein with the highest degree of similarity to NMP1 in the fully sequenced yeast genome (28% identity; 55% similarity). The next closest \(S.\ cerevisiae\) relative to NMP1 is tropomyosin, another filament-like cytoskeletal protein (26% identity, 51% similarity). In the \(C.\ elegans\) genome, no match was found with more than 19% identity represented by an uncharacterized open reading frame.

By contrast to the few weak similarities to animal and fungal proteins, a number of ESTs encoding potential orthologues of NMP1 were identified from different plant species, including the liverwort \(Marchantia polymorpha\) and the moss \(Physcomitrella patens\) (Fig. 2B). The amino acid sequences of these orthologues are highly conserved. For example, the tomato and maize full-length sequences share 69% identical residues, with many of the non-matching residues being conservative exchanges (Fig. 2B). The 42 aa open reading frame of the \(Marchantia\ polymorpha\) EST is 74% identical to the tomato NMP1 protein (LeNMP1). Similar degrees of conservation were found with sequences from the dicots \(Arabidopsis thaliana\), bean, and soybean, the monocots wheat, rye and lotus, the gymnosperm \(Cryptomeria japonica\) (Japanese cedar), and the moss \(Physcomitrella patens\), with the regions of similarity spanning the entire NMP1 sequence (Fig. 2B).

In order to characterize NMP1, the subcellular localization of the protein was investigated. An antiserum raised against recombinant LeNMP1 was used to detect NMP1 in plant protein extracts (Fig. 3). The antiserum detects a protein of c. 36 kDa in tobacco leaves (Fig. 3A, lane 2) comparable to the tomato leaf extract loaded as a control (Fig. 3A, lane 1), as well as in tobacco BY-2 suspension culture cells (Fig. 3A, lane 3) and tobacco NT-1 suspension culture cells (data not shown). A smaller amount of NMP1 is present in isolated BY-2 nuclei (Fig. 3A, lane 4). Significantly, in a core nuclear matrix fraction from which more than 90% of the histones are removed (Fig. 3B, compare lanes 4 and 5) NMP1 is retained (Fig. 3A, compare lanes 4 and 5). The increase in NMP1 signal intensity on the protein blot correlates with the about 4-fold increase in numbers of organelles (c. 250 000 nuclei in lane 4, c. 1 000 000 nuclear matrices in lane 5) used to achieve approximately equal loading of the non-histone nuclear proteins as shown by the Coomassie staining. This standard assay for the identification of nuclear matrix proteins indicates that the nuclear-localized fraction of NMP1 is associated with the operationally defined nuclear matrix. No signal was detected when probing an equal amount of rat nuclear matrix proteins with the anti-NMP1 antibody (data not shown).

To characterize the subcellular localization of NMP1 further, the antiserum was used for immunofluorescence microscopy with tobacco NT-1 protoplasts. The confocal optical sections shown in Fig. 4A, B, and C through the centre of the nucleus confirm that NMP1 is localized both in the nucleus and in the cytoplasm. The nuclear fraction of the protein is excluded from the nucleolus. No signal was detected with the preimmune serum under identical conditions (Fig. 4D, E, F). An identical localization pattern was detected \(in\ vivo\) for an NMP1-GFP fusion protein in transiently transformed BY-2 cells (Fig. 4G–J). The 64 kDa NMP1-GFP fusion protein is present both inside the nucleus and in the cytoplasm (Fig. 4G, H). A fusion protein of comparable size consisting of a non-targetable deletion mutant of \(Arabidopsis\) RanGAP fused to GFP (AtRanGAP1ΔN-GFP; Rose and Meier, 2000) is clearly excluded from the nucleus (Fig. 4I, J). This demonstrates that a fusion protein of the size of NMP1-GFP does not passively diffuse into the nucleus. Together, these data show that NMP1 is localized both in the cytoplasm and the nucleus of plant cells and that the nuclear fraction of NMP1 is associated with the operationally defined insoluble nuclear matrix.

No indication for the presence of nuclear localization signals could be found analysing the sequences of LeNMP1 (accession number AAK83083) and AtNMP1 (accession number AAM64307) using the programs
PredictNLS (http://maple.bioc.columbia.edu/predictNLS/), Motif Scan (http://hits.isb-sib.ch/cgi-bin/PFSCAN), and PSORT (http://psort.nibb.ac.jp/). It is well known, however, that several animal structural nuclear proteins do not contain an NLS. The precise mechanism of their import is presently not known.

An RNA blot analysis showed that the 1.35 kb NMP1 transcript is present in leaves, young fruit, flowers, and light- and dark-grown seedlings of tomato. Minor differences in mRNA abundance were observed between tissues (Fig. 5A). An immunoblot with total protein extracts from leaves, young fruit, flowers, seedlings, red fruit, and roots showed the presence of a c. 36 kDa protein in all organs (Fig. 5B). Again, small differences in abundance were detected, but the overall expression pattern of both mRNA and protein indicates that NMP1 is a ubiquitous protein in tomato.

Discussion

NMP1 is a novel, highly α-helical plant protein, which has a low degree of similarity to other α-helical and filament-like proteins from animals and fungi. However, the 74% amino acid identity and 95% similarity between LeNMP1 and a 42 aa open reading frame from the liverwort Marchantia polymorpha indicate a striking degree of sequence conservation throughout the plant kingdom. This high degree of conservation of the NMP1 protein sequence among flowering and non-flowering plants is in stark contrast to the lack of similar sequences in the complete genomes of yeast and C. elegans, and indicates a plant-specific function of NMP1.

The most significant structural motif on NMP1 is a large number of short heptad-repeats with hydrophobic amino acids at positions $a$ and $d$, typical for proteins forming coiled-coil protein-protein interactions (Lupas et al., 1991). Comparison of the plant NMP1 sequences showed
that all hydrophobic residues predicted to be involved in forming coiled-coil domains, except for two residues in *Marchantia polymorpha*, are conserved between all sequences, indicating a functional importance of this predicted tertiary structure of NMP1. The repeats are in the range of those functionally verified in other proteins and shown to form coiled-coil structures (Barthe et al., 2000; Pelletier et al., 1997; Shotland et al., 2000).

NMP1 is localized both in the cytoplasm and the nucleus of tobacco cells, and appears to be excluded from the nucleolus. The retention of NMP1 in the nuclear matrix fraction of BY-2 cells, from which more than 90% of the histones are removed, indicates that the nuclear portion of NMP1 is tightly associated with the operationally defined nuclear matrix. The absence of an immunoreactive protein in rat nuclear matrix suggests that NMP1 might be a plant-specific protein and conforms with the lack of homologous sequences outside the plant kingdom.

NMP1 does not appear to contain a classic bipartite or monopartite nuclear localization signal. The size of NMP1 (37.4 kDa) is slightly below the experimentally determined exclusion size of the nuclear pore (c. 40 kDa; Grebenok et al., 1997), which might allow the native protein to diffuse passively into the nucleus. However, the NMP1-GFP fusion protein, which is significantly larger than 40 kDa, also enters the nucleus. There have been several reports recently of proteins passing the nuclear pore without the presence of a ‘classical’ NLS (Whitehurst et al., 2002; Miyamoto et al., 2002). In addition, evidence exists that several structural nuclear proteins enter the nucleus in the absence of an NLS (Wu et al., 2002). Further studies will be required to determine by which mechanism NMP1 enters the plant nucleus.

The majority of NMP1 is localized in the cytoplasm. All nuclear matrix proteins found to date in animal cells are located either exclusively (Nakayasu and Berezney, 1991; Belgrader et al., 1991) or predominantly (Mortillaro and Berezney, 1998) in the cell nucleus, based on both subfractionation by immunoblots and immunofluorescence analysis. The fractionation pattern of NMP1 is clearly different from these animal examples. There are some examples known from animals of cytoplasmic cytoskeletal proteins also involved in nuclear functions (Simcha et al., 1998; Petit et al., 2000). Presently, nothing is known about plant cytoskeletal proteins also located in the nucleus. Its association with the nuclear matrix might indicate that it is involved in an architectural role at least in the nuclear compartment.

Acknowledgements

We would like to thank Dr Ronald Berezney (State University of New York at Buffalo) for a rat nuclear matrix preparation, and Nancy Peffer (Dupont Pharmaceuticals) for excellent help with the protein expression and purification for antiserum production. This work was supported in part by a grant from the National Science Foundation (MCB-0079577).

References


Whitehurst AW, Wilsbacher JL, You Y, Luby-Phelps K, Moore MS, Cobb MH. 2002. ERK2 enters the nucleus by a carrier-

