Plant Molecular Biology **19**: 847–857, 1992.
© 1992 *Kluwer Academic Publishers. Printed in Belgium.*

Molecular characterization of tobacco cDNAs encoding two small GTP-binding proteins

Geza Dallmann [1,3], Liliane Sticher [1], Christopher Marshallsay [1] and Ferenc Nagy [1,2]*
[1] *Friedrich-Miescher Institute, P.O. Box 2543, CH-4002 Basel, Switzerland (* author for correspondence);* [2] *Biological Research Center, Hungarian Academy of Sciences, P.O. Box 521, H-6701 Szeged, Hungary;* [3] *Present address: Agricultural Biotechnology Center, P.O. Box 170, H-2100 Godollo, Hungary*

Received 20 January 1992; accepted in revised form 10 April 1992

Key words: GTP-binding proteins, tobacco, *ras*-related genes, tissue specificity

Abstract

We have isolated two cDNAs encoding small GTP-binding proteins from leaf cDNA libraries. These cDNAs encode distinct proteins which show considerable homology to members of the *ras* superfamily. *Np-ypt3*, a 1044 bp long *Nicotiana plumbaginifolia* cDNA, encodes a 24.4 kDa protein which shows 65% amino acid sequence similarity to the *Schizosaccharomyces pombe* ypt3 protein. The *N-ypt3* gene is differentially expressed in mature flowering plants. Expression of this gene is weak in leaves, higher in stems and roots, but highest in petals, stigmas and stamens. *Nt-rab5*, a 712 bp long *Nicotiana tabacum* SR1 cDNA, encodes a 21.9 kDa protein which displays 65% amino acid sequence similarity to mammalian *rab5* proteins. The expression pattern of the *Nt-rab5* gene is very similar to that of the *Np-ypt3* gene. The *Nt-rab5* gene is virtually not expressed in leaves, higher in stems and roots, and highest in flowers. Both the *Nt-rab5* and *Np-ypt3* proteins were expressed in *Escherichia coli* and shown to bind GTP.

Introduction

In eukaryotic cells a variety of small (21–25 kDa) GTP-binding proteins have been described that are thought to regulate diverse cellular responses. Based on sequence analysis these proteins can be conveniently divided into three groups: *ras* and its close relatives [5], *rho* and its close relatives [10], and the *rab* family [36, 37]. The *rab* family is considered to be the human counterpart of the yeast *ypt/sec4* proteins. Since the discovery of the *ypt1* gene in *Saccharomyces cerevisiae* [11] more than 30 genes encoding small GTP-binding proteins have been cloned from mammals, *Drosophila*, and yeast (reviewed in Hall [15]). These genes encode proteins ca. 200 amino acids long which display significant homology, especially within the regions involved in GTP binding and GTP hydrolysis.

The precise physiological functions of proteins encoded by *ras* and *ras*-related genes are not fully understood. However, increasing evidence suggests that the prototype for this family, the *ras* protein, and its close relatives (*ral* and *rab* proteins) may be involved in signal transduction pathways [5, 15]. Recent reports indicate that the *ypt/sec* proteins are regulatory elements in the secretory pathway of *S. cerevisiae* (reviewed in [4]). The *ypt1* protein, which is presumed to be associated with Golgi-like structures [32], is an

essential component for the vesicular transport of proteins from the endoplasmic reticulum to the Golgi apparatus [3]. The *sec4* protein was found to be associated with the cytoplasmic face of the plasma membrane and with post-Golgi-like secretory vesicles. Biochemical evidence shows that it is involved in regulating protein transport between the Golgi apparatus and the plasma membrane [30].

By analogy to these yeast proteins, it has been proposed that the mammalian *rab* proteins are also specific regulators of membrane traffic. This view is consistent with the observation that individual *rab* proteins are located in different intracellular compartments along the endocytic and exocytic pathways [13, 8]. Moreover, the hypervariable C-terminal domain of *rab* proteins was shown to act as signals for the proper association of *rab* proteins with their specific target membranes in the endocytic pathway [9].

Despite the essential physiological roles of these proteins in yeast and mammals, little is known about small GTP-binding proteins in higher plants. Only a few plant GTP-binding proteins [6, 16] or genes encoding GTP-binding proteins have been identified. Recently *ras*-related genes from *Arabidopsis thaliana* [2, 22], *Zea mays* [29], and *Oryza sativa* [31] were reported. Here we report the characterization of two tobacco cDNAs encoding distinct small GTP-binding proteins. These cDNAs, *Nt-rab5* and *Np-ypt3*, encode proteins very similar in sequence to the mammalian *rab5* and fission yeast *ypt3* proteins. The genes *Nt-rab5* and *Np-ypt3* exhibit a nearly identical expression pattern. They are expressed predominantly in petals, stamens, stigmas and show only a very low level of expression in leaves and in young developing tobacco seedlings.

Materials and methods

Design and synthesis of oligonucleotides for polymerase chain reactions (PCR)

Small GTP-binding proteins from diverse organisms share highly conserved sequences, including the motifs WDTAGQ and FMETSA [22, 29]. We have chemically synthesized two degenerate 17-mer oligonucleotides corresponding to these sequence motifs by using an Applied Biosystems 380A DNA Synthesizer. Oligo A, sequence 5′-TGGGA[TC]AC[TCGA]GC[TCGA]GG[TCGA]CA, encodes the peptide sequence WDTAGQ. Oligo B, sequence 5′-GC[TCGA]GA[TCGA]GT[TC]TCCATAAA, corresponds to the peptide sequence FMETSA but in the antisense orientation.

Preparation of cDNAs, PCR assays and cDNA libraries

Poly(A)$^+$ RNA, isolated from fully developed laves of *N. tabacum* SR1 and *N. plumbaginifolia* plants grown in the greenhouse under 16L/8D conditions, was used to prepare cDNA using the Amersham cDNA Synthesis Plus kit and protocols. Ca. 100 ng of this cDNA was then used in PCR amplifications, using Cetus protocols with oligos A and B, for 30 cycles with an annealing temperature of 50 °C.

The *N. plumbaginifolia* leaf cDNA library constructed in λgt10 was a generous gift from Dr M. Mieszczak. The *N. tabacum* SR1 leaf cDNA library, constructed in λ-ZAPII, was purchased from Stratagene.

Screening of cDNA libraries, cloning and recombinant DNA techniques

All DNA manipulations were carried out as described [21]. 2.5×10^8 recombinant phage from each of the two cDNA libraries described above were screened by plaque hybridization using DNA fragments, encoding selected, PCR-amplified sequences, labelled with [α-^{32}P]ATP by random priming, as probes. Filters were prehybridized for 4 h at 50 °C in 6 × SSC, 5 × Denhardt's solution, 0.1% SDS, 100 μg/ml denatured salmon sperm DNA. Hybridization was performed in the same solution supplemented with 5×10^7 cpm of probe for 20 h at 52 °C. Filters

were washed with 6 × SSC and 0.1% SDS for 1 h at room temperature and twice for 10 min at 60 °C with 0.2 × SSC, 0.1% SDS. After washing, filters were briefly rinsed in sterile H_2O, air-dried and autoradiographed. Positive phage were isolated and their respective cDNA inserts then subcloned into either the *Eco* RI or *Not* I site of the vector pKS and sequenced using the dideoxy chain termination method.

RNA isolation and northern hybridization

Total and poly(A)$^+$ RNA was isolated from different tissues of mature flowering tobacco plants grown in a greenhouse under 16L/8D cycles. Alternatively, we isolated RNA from 2-week-old tobacco seedlings grown on MS medium without phytohormones or from callus maintained on MS medium supplemented with 1.0 μg/ml NAA and 0.1 μg/ml BAP. Total and poly(A)$^+$ RNA was isolated as described [27]. For northern hybridizations, 20 μg of total RNA and/or 2 μg of leaf poly(A)$^+$ RNA were denatured and fractionated on a 1.5% formaldehyde-agarose gel. Probes were labelled with [α-^{32}P]ATP by random priming and purified on a Sephadex G-50 column. Gels were blotted onto nylon membrane filter (Hybond-N, Amersham). After crossing with UV light, the filters were prehybridized and hybridized as described for plaque hybridization. Washing of the filters hybridized with the full-length cDNA or with fragments containing the 3′-untranslated regions of the corresponding genomic clones were, however, performed under more stringent conditions. Northern-blot filters were washed for 1 h at room temperature and subsequently twice at 65 °C with 0.1 × SSC, 0.1% SDS.

Expression of proteins in Escherichia coli

The GST Gene Fusion System (Pharmacia) [35] was used to express the proteins encoded by the plant cDNAs as fusion proteins in *E. coli*. For convenient cloning we modified the plant cDNAs by using PCR with synthetic oligonucleotides which introduced *Bam* HI sites at the start and at the stop codons. The modified, PCR-amplified cDNA sequences were then cut with *Bam* HI and cloned into the appropriately cut pGEX-2T vector. Expression of these constructs would result in the protein of interest fused to the carboxy terminus of glutathione *S*-transferase (GST). Clones containing the plant cDNAs in the sense and, as controls, antisense orientation were then transferred into HB101 cells. Growth, induction, and preparation of cell extracts were performed as described [35]. The expressed GST plant fusion proteins were then purified from the total crude bacterial extracts by repeated glutathione Sepharose 4B column chromatography. The purified products were then finally analysed by 15% SDS-PAGE as described [19].

GTP-binding assays

For GTP-binding assays either total crude bacterial extracts isolated from induced and uninduced cells or purified fusion proteins were used. Aliquots of different extracts were transferred to a BA 85 nitrocellulose filter by using a Schleicher & Schuell SRC 96D Minifold I Dot Blotter. Filters were then incubated with [α-^{32}P]GTP and washed as described [32]. After washing the filters were dried and autoradiographed for 10 h at −70 °C.

Results

Isolation of PCR products with homology to ras-*related proteins*

We have used a PCR approach to identify plant cDNAs which encode small GTP-binding proteins. Leaf cDNA was prepared and used as template in PCR amplifications with degenerate oligonucleotides corresponding to the highly conserved WDTAGQ and FMETSA motifs, found in members of the *ras* superfamily, as primers. The PCR-amplified DNA fragments obtained, ca. 270 bp long, were cloned and

sequenced. Computer databank searches were then carried out with the deduced amino acid sequences. Sequences which exhibited more than 50% homology to known *ras*-related proteins or to specific *ras*-related sequence motifs were then kept. We have identified so far more than 15 different sequences which show considerable homology to sequence motifs present in *ras* or *ras*-related proteins. We are presently screening cDNA libraries with these characterized fragments to identify full-length cDNAs encoding these putative small GTP-binding proteins. Screening with two such PCR-amplified sequences, designated *PCR-rabA* and *PCR-yptA*, has already resulted in the isolation and characterization of several cDNA clones belonging to two distinct groups: (i) cDNAs encoding proteins very similar to the mammalian *rab5* proteins, and (ii) cDNAs encoding proteins with considerable homology to the *ypt3* protein of *S. pombe*.

Characterization of Nt-rab5 *cDNA*

Tobacco is an amphidiploid species believed to be derived from ancestral species most closely related to present-day *N. sylvestris* and *N. tomentosiformis*. As a consequence, several tobacco gene

```
  1  CTGTCTCTTCCATTAATTTCGTCTCTTAATAGCTTACAACTCAATATCTCAGAAACACAC  60

 61  TGTAAATTCTTCATTTTTACCTCATCAGTCATTTGATCATGGCGTCAAGGCGTCACAATA  120
                                                M  A  S  R  R  H  N  N

121  ATCTCAATGCAAAACTCGTGTTATTAGGGGACATGGGAGCTGGTAAATCAAGCTTGGTTA  180
       L  N  A  K  L  V  L  L  G  D  M  G  A  G  K  S  S  L  V  I

181  TACGATTCGTCAAGGGTCAATTCCTTGAATTCCAGGAATCGACGATCGGAGCGGCGTTCT  240
       R  F  V  K  G  Q  F  L  E  F  Q  E  S  T  I  G  A  A  F  F

241  TTTCGTCAACGGTGTCAGTTAACAATGCAACGGTGAAGTTTGAGATCTGGGATACTGCTG  300
       S  S  T  V  S  V  N  N  A  T  V  K  F  E  I  W  D  T  A  G

301  GTCAGGAGAGGTACCACAGCTTAGCGCCTATGTACTACAGAGGTGCTGCAGCTGCTATCA  360
       Q  E  R  Y  H  S  L  A  P  M  Y  Y  R  G  A  A  A  A  I  I

361  TCGTGTATGACATCACAAGCACTGAATCACTTGCACGAGCAAAGAAATGGGTTCAAGAGT  420
       V  Y  D  I  T  S  T  E  S  L  A  R  A  K  K  W  V  Q  E  L

421  TGCAGAAGCAAGGTAATCCCAACATGGTCATGGCTCTTGCTGGAAACAAAGCCGATCTAG  480
       Q  K  Q  G  N  P  N  M  V  M  A  L  A  G  N  K  A  D  L  E

481  AAGATAAGAGGAAGGTGACTGCAGAAGAAGCACGTCTGTACGCAGAGGAAAACGGTCTTT  540
       D  K  R  K  V  T  A  E  E  A  R  L  Y  A  E  E  N  G  L  F

541  TTTTCATGGAAACCTCTGCCAAAACTGCTACCAATGTCAATGATATTTTCTACGAAATAG  600
       F  M  E  T  S  A  K  T  A  T  N  V  N  D  I  F  Y  E  I  A

601  CTAAACGGTTGCCTAGAGCTCAACCTGCACAAAATCCGGCAGGAATGGTACTAGAAGACA  660
       K  R  L  P  R  A  Q  P  A  Q  N  P  A  G  M  V  L  E  D  K

661  AACCAGCGCAAGGATCACAAGCTGCATCATGCTGTACTTAAGTATTATCATG  712
       P  A  Q  G  S  Q  A  A  S  C  C  T  *
```

Fig. 1. Nucleotide and deduced amino acid sequence of the *Nt-rab5* cDNA. The amino acid sequence is indicated by the conventional single-letter code. Degenerate oligonucleotides used in PCR correspond to the underlined regions.

families studied in detail consist of subfamilies from both origins [14]. If genes of the two ancestral lines are uniformly expressed, one would expect that isolated cDNA-derived PCR sequences will be derived from genes which are closely related but not identical. Indeed, we have identified six PCR-derived sequences (each ca. 270 bp long), which are very similar to the mammalian *rab5* sequences, and belong to two distinct groups which show more than 95% identity at nucleotide sequence level (data not shown). Mismatches were found only at the 'wobble position' and therefore these PCR-derived sequences probably represent two very closely related genes. Full-length cDNAs were identified using the *PCR-rabA* fragment as a probe for cDNA library screening. We have isolated 6 cDNA clones and sequenced the longest cDNA insert. This cDNA clone, designated *Nt-rab5*, contains a 712 bp insert with a 600 bp ORF which would encode a 200 amino acid long polypeptide with a molecular mass of 21.9 kDa (Fig. 1). The putative start codon is placed at position 100 and the stop codon at position 700. The putative start codon is preceded by three in-frame stop codons at positions 27, 30 and 63. The deduced amino acid sequence of the *Nt-rab5* protein showed 65% and 61% identity with the human and canine *rab5* proteins, respectively. Moreover, the *Nt-rab5* protein also shows 30–40% homology with other members of the *ras* superfamily (data not shown). The *Nt-rab5* protein displayed the highest homology to the recently described *A. thaliana Rha-1* protein. These two plant proteins show 78% amino acid sequence identity (87% similarity) and diverge only towards the C-terminus (Fig. 2). An amino acid sequence alignment of *Nt-rab5* protein with the *A. thaliana Rha-1* and human *rab5* proteins is shown in Fig. 2. This comparison clearly demonstrates that regions shown to participate in the formation of the GTP-binding site [28] are well conserved among these proteins. Indeed, the G-X-X-X-X-G-K-S (positions 28–35) and D-X-X-G (positions 76–69) protein motifs involved in the binding of the phosphoryl residues of GTP as well as the N-K-X-D (positions 134–137) and E-T-S-A (162–165) motifs considered to be in-

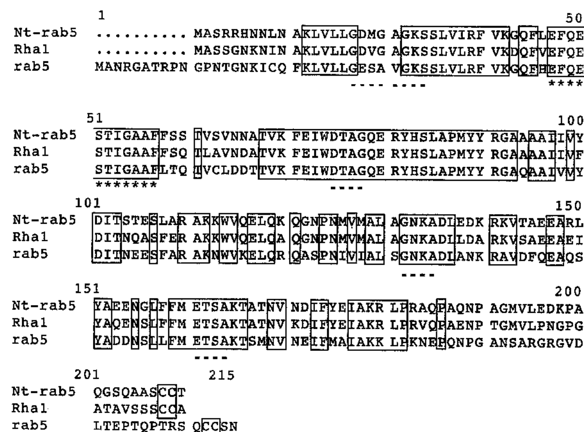

Fig. 2. Comparison of the deduced amino acid sequences of the *Rha-1* and *rab5* gene products with those of *Nt-rab5*. Sequences are from *A. thaliana Rha-1* [2], canine *rab5* [8] and tobacco *Nt-rab5*. Amino acids residues conserved in all three proteins are boxed. The putative effector region is indicated by asterisks. Regions implicated in GTP-binding are underlined by dashed lines. Numbering refers to residues in *Rha-1*.

volved in the binding of the guanine moiety are nearly identical in these three proteins. Moreover, the two cysteine residues as the C-terminus shown to be essential for membrane attachment are also conserved [20]. Finally, the 'effector binding' domain E-F-Q-E-S-T-I-G-A-A-F [34] between positions 47–55 which is thought to interact with other proteins such as the GAP-related proteins [1] is identical in these proteins. Based on the amino acido sequence similarity over their full length and also because of the identical effector domains, we suggest that the *Nt-rab5* and *Rha-1* proteins are the functional counterparts of the mammalian *rab5* protein.

Characterization of the Np-ypt3 *cDNA*

Among the isolated, sequenced, PCR-amplified products obtained from the *N. plumbaginifolia* leaf cDNA, we found 5 DNA sequences whose deduced amino acid sequences are very similar to that of the *S. pombe ypt3* protein. These products were completely identical over their 270 bp length and showed more than 98% homology to PCR products isolated by using *N. tabacum* leaf cDNA. To isolate a full-length cDNA clone we have

```
   1  TTGAAGCAGAGTTCTGTGATTTACTTACTTTTTTGCCTAACCAAAAAACACAGAACAAAG   60

  61  TAAGTATTTTCAACGATGGCAGGTTATAGAGCAGATGATGAGTACGATTACTTATTCAAG  120
                     M  A  G  Y  R  A  D  D  E  Y  D  Y  L  F  K

 121  CTAGTGTTGATCGGAGATTCAGGTGTGGGTAAATCAAATCTGCTTTCAAGGTTCACTAAA  180
      L  V  L  I  G  D  S  G  V  G  K  S  N  L  L  S  R  F  T  K

 181  AATGAATTCAATTTGGAGTCTAAGTCTACAATTGGTGTTGAGTTTGCTACCAAAAGTCTC  240
      N  E  F  N  L  E  S  K  S  T  I  G  V  E  F  A  T  K  S  L

 241  AATATTGATAATAAAGTTATTAAAGCTCAGATTTGGGATACTGCTGGCCAAGAAAGATAT  300
      N  I  D  N  K  V  I  K  A  Q  I  W  D  T  A  G  Q  E  R  Y

 301  CGTGCCATTACTAGTGCTTATTATCGGGGAGCTGTTGGCGCTTTGCTCGTATATGATGTA  360
      R  A  I  T  S  A  Y  Y  R  G  A  V  G  A  L  L  V  Y  D  V

 361  ACTCGACATGTTACCTATGAGAACGTCACAAGATGGTTGAAGGAATTGAGAGACCATACT  420
      T  R  H  V  T  Y  E  N  V  T  R  W  L  K  E  L  R  D  H  T

 421  GACCCTAACATTGTAGTTATGCTCATAGGCAACAAGTCAGATCTCCGCCATCTCGTTGCT  480
      D  P  N  I  V  V  M  L  I  G  N  K  S  D  L  R  H  L  V  A

 481  GTTTCAACTGATGAGGCTAAAGGTTTGGCAGAGAGGGAGGGCCTATACTTTATGGAGACT  540
      V  S  T  D  E  A  K  G  L  A  E  R  E  G  L  Y  F  M  E  T

 541  TCTGCATTAGAAGCAACCAACGTGGAAAATGCATTCACTGAAGCTCTCACGCAGATATAC  600
      S  A  L  E  A  T  N  V  E  N  A  F  T  E  A  L  T  Q  I  Y

 601  CGGATCGTCAGTAAAAAAGCAGTCGAGGCAGGCGATGAAGGTGCTACTTCATCTGCTCCT  660
      R  I  V  S  K  K  A  V  E  A  G  D  E  G  A  T  S  S  A  P

 661  CCTAAAGGAGAGACGATTAACATCAAAGATGAAGGGTCTTCCTGGAAAAAGTTCGGATGC  720
      P  K  G  E  T  I  N  I  K  D  E  G  S  S  W  K  K  F  G  C

 721  TGTTCAAGCTAGGGCTCTTGAAAAATGTGCAAGACAGATGCTCAAGCTGTTTTAGTTGTA  780
      C  S  S  *        *

 781  AATTGTCATTAAATCAATGATTTCCTCCCCCAGTATTCCTTATCCTTTTTCGGTTATCTT  840

 841  GGTTGTGGGATAAGATTTAAGCTAGCCATGAGAATACACGAAGCAGCCTTCTCATGAAAA  900

 901  ACTCCGAAGGAGCAGGTACCGTTGGTTTCTAATTACAGAATGTGCAAACACGGTCTTAGT  960

 961  TTTATGTGTTCTGGATGTTAGCTTGAAAAAATGGACAGTGGCAAGATCTTTCCAATAACA 1020

1021  TGTTACCGGCGGCCGCTCTAGAAC  1044
```

Fig. 3. The nucleotide sequence of he *Np-ypt3* cDNA. The deduced amino acid sequence is given by using the conventional single-letter code. Degenerate oligonucleotides used in PCR correspond to the underlined sequences.

screened a *N. plumbaginifolia* leaf cDNA library, using one of the five identical PCR fragments, *PCR-yptA*, as a probe. The longest cDNA clone isolated (1044 bp long), designated *Np-ypt3*, was sequenced and found to encode a 218 amino acid long polypeptide, with a molecular mass of

24.4 kDa (Fig. 3). The putative start codon is placed at position 75 and the two stop codons at 730 and 740, respectively. The putative start codon is preceded with an in-frame stop codon at position 61. The protein encoded by this cDNA displays 65% similarity to the *S. pombe ypt3* protein [24], 56% similarity to the protein encoded by the *ras*-related *ara* gene of *A. thaliana* [22], 57% similarity to the *Zea mays yptm*-encoded protein [29], and 64% homology to the protein encoded by the *rgp1* gene of *O. sativa* [31]. The *Np-ypt3* protein also exibits amino acid sequence similarity to other *ras*-related proteins such as *ypt2* (40%) [17] and *H-rab* (47%) [37]. The observed homology suggested that the *Np-ypt3* protein belongs to the *ypt* family. An alignment of the sequences most highly conserved between *Np-ypt3*, *S. pombe ypt3*, and *O. sativa rgp1* proteins is shown in Fig. 4. Regions of these proteins which may be involved in GTP binding and hydrolysis are well conserved. These regions, G-X-X-X-X-G-K-S (positions 25–32), D-X-X-G (positions 73–76) as well as the N-K-X-D (positions 140–144) and E-T-S-A motifs (positions 159–162) are nearly identical in the three proteins. In contrast, the N- and C-terminal of these proteins are very different from each other. However, as in other *ras*-related proteins the two C-terminal cysteine residues essential for membrane attachment are conserved. The two cysteine residues are followed by SS and MS motifs in the plant protein. Finally, the 'effector binding' domain of these proteins, S-K-S-T-I-G-V-E-F between positions 47–54, also shows striking similarity, suggesting that these proteins interact with identical or very similar GAP proteins. The similarity between the *O. sativa rgp1* and the *Np-ypt3* proteins (65%) is lower than the similarity of these plant proteins to the yeast *ypt3* protein. This binding suggests that in spite of their similarity these plant proteins are not interchangeable and that their physiological roles in the cell could differ.

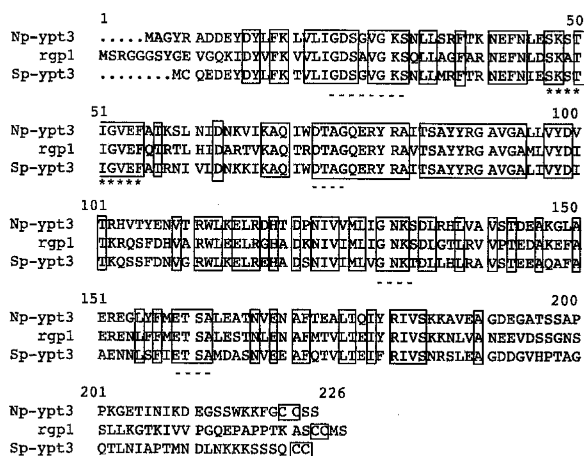

Fig. 4. Comparison of the deduced amino acid sequences of the *rgp1* and *ypt3* genes to those of *Np-ypt3*. Proteins are from *O. sativa rgp1* [31], *S. pombe ypt3* [24] and *N. plumbaginifolia Np-ypt3*. The identical amino acid residues conserved in all three proteins are boxed. The putative effector binding domain is indicated by asterisks. Regions implicated in GTP-binding are underlined by dashed lines. The numbering refers to the sequence of the *rgp1* protein.

To identify possible physiological functions of *Nt-rab5* and *Np-ypt3* proteins, we have measured their mRNA expression levels in different tissues by northern blot analysis. Total and poly(A)$^+$ RNA isolated from different tissues of mature, flowering tobacco plants were hybridized with full-length cDNAs or with fragments containing the 3′-untranslated regions of the corresponding genomic clones. (Sequences of the genomic clones will be published elsewhere.) Figures 5A and 5C show that both genes hybridized to transcripts of approximately 1.2 kb. The levels of *Nt-rab5* and *Np-ypt3* mRNAs were consistently high in petals, anthers and stigmas, and were expressed only at a very low level in fully developed leaves and in young, two-week-old developing seedlings. The abundance of *Nt-rab5* and *Np-ypt3* in mRNA from callus cells was also low. Moreover, the expression of these mRNAs was not affected by light. We found that expression patterns by using the full-length cDNAs or the corresponding 3′-untranslated regions as probes were identical in both cases. The differential levels of *Np-ytp3* and *Nt-rab5* were, however, clearly mRNA-specific since levels of *Cab* mRNA, encoding the chlorophyll *a*/*b*-binding proteins, showed different patterns (Fig. 5B). Expression of the *Cab* genes was high in leaves, 10-fold lower in stems and stigmas, and below detection levels in other tissues. Moreover, consistent with previously reported data the

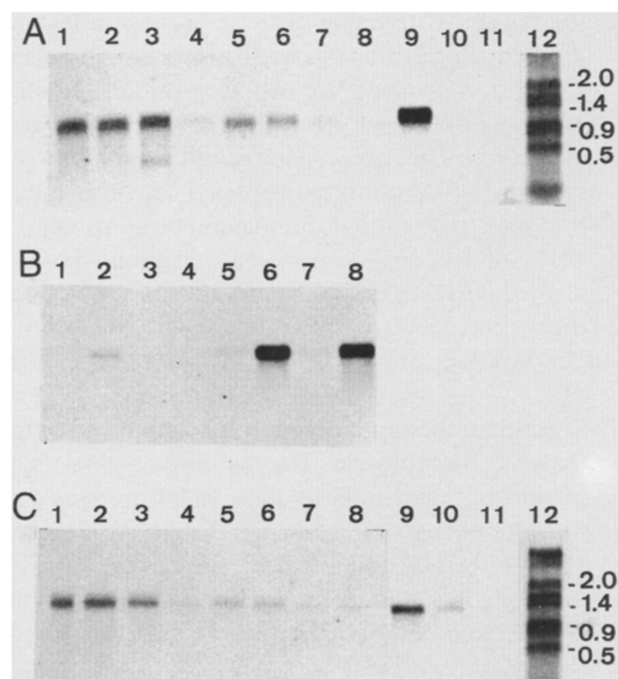

Fig. 5. Northern hybridization analysis of the *Nt-rab5* and *Np-ypt3* gene expression. The hybridization probes were (A) full-length *Nt-rab5* cDNA, (B) full-length *Cab1* cDNA and (C) full-length *Np-ypt3* cDNA. Probes were labelled with a random primer kit. Total RNA samples are from stamens (lane 1), petals (lane 2), stigmas (lane 3), light-grown leaves (lanes 4, 7), stems (lane 5), roots (lane 6), dark-adapted leaves (lane 8), callus (lane 10), and 14-day-old seedlings (lane 11) of *N. tabacum* SR1 (A) and of *N. plumbaginifolia* (C). Each lane, except lane 9, contains 20 μg of total RNA. Lane 9 contains 2 μg leaf poly(A)$^+$. Lane 12 contains size marker (kb): labelled λ-DNA *Hind* III-*Eco* RI digest. B. Each lane contains 20 μg total RNA from stamens (lane 1), stigmas (lane 2), petals (lane 3), roots (lane 4), stems (lane 5), light-grown leaves (lane 6), dark-adapted leaves (lane 7), and light-grown 24-day-old tobacco seedlings (lane 8) of tobacco. RNAs used in these experiments are from the same preparations.

Cab genes were expressed in a light-dependent manner [26].

GTP-binding activity of Nt-rab5 *and* Np-ypt3 *proteins*

To verify that the *Nt-rab5* and *Np-ypt3* proteins are indeed GTP-binding proteins we have carried out *in vitro* GTP-binding assays. To obtain large quantities of these proteins we have expressed their cDNAs in *E. coli* using the GST gene fusion system. The fusion proteins obtained were purified and then transferred to nitrocellulose filters and incubated with [α-^{32}P]GTP according to Schmitt *et al.* [32]. Autoradiography showed that GST fusion proteins with either the *Np-ypt3* or the *Nt-rab5* proteins bind [α-^{32}P]GTP (Fig. 6, panels 3–6). No GTP-binding activity is shown in control experiments with wild-type GST (panels 1, 11), bovine serum albumin (panel 2), *E. coli* extracts containing control GST hybrid proteins (panels 7–10), or *E. coli* extracts not expressing GST of hybrid proteins (panel 12). The observed GTP-binding is therefore specific to the *Np-ypt3* and *Nt-rab5* hybrid proteins.

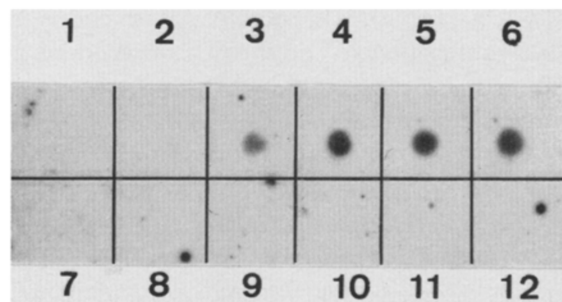

Fig. 6. GTP-binding analysis of the *Nt-rab5* and *Np-ypt3* proteins. Samples are (1) 100 μg glutathione *S*-transferase (GST), (2) 100 μg BSA, (3) total crude extract (100 μg protein) from induced *E. coli* cells containing the GST/*Nt-rab5* fusion protein (cDNA expressed in the sense orientation), (4) 10 μg purified GST/*Nt-rab5* fusion protein, (5) total crude extract (100 μg protein) from induced cells containing the GST/*Np-ypt3* protein (cDNA expressed in the sense orientation), (6) 10 μg purified GST/*Np-ypt3* fusion protein. Total crude extracts (100 μg in each case) from induced and uninduced cells containing (7 & 8) the *Nt-rab5* (cDNA in the antisense orientation), or (9 & 10) the *Np-yt3* (cDNA in the antisense orientation). Total crude extracts (100 μg in each) from cells containing the vector pGEX-2T, (11) after induction, and (12) uninduced.

Discussion

In this paper we describe the isolation and characterization of two plant cDNAs, designated *Np-ypt3* and *Nt-rab5*. We have shown that these cDNAs encode distinct *ras*-related small GTP-binding proteins. The *Np-ypt3* and *Nt-rab5*

cDNAs encode polypeptides containing 218 and 210 amino acids respectively. These cDNAs were expressed in *E. coli* and produced proteins which exhibit GTP-binding activities. The predicted amino acid sequence of the *Np-ypt3* gene is most closely related to the *S. pombe ypt3* protein [24] and to the recently described *O. sativa rgp1* protein [31]. The *Nt-rab5* protein shows the highest sequence homology to the mammalian *rab5* proteins and the *A. thaliana Rha-1* protein [2]. Conserved regions believed to be implicated in GTP-binding and hydrolysis were highly conserved in both the *Nt-rab5* and *Np-ypt3* proteins.

Southern hybridization analysis (data not shown), under stringent conditions, indicates the presence of one *rab5* and *ypt3* type gene per haploid genome in *N. plumbaginifolia*. Hybridization analysis under increasingly moderate conditions finally revealed about ten hybridizing bands. It is likely, however, that some of the hybridizing bands (because of the low stringency conditions) will represent members of other subfamilies of the *ras* superfamily present in higher plants. These observations, together with the large number of PCR-generated *ras*-related sequences, clearly indicate a high complexity of genes encoding small GTP-binding proteins, but the exact number of the *ypt/rab* type genes remains to be seen.

So far, including the two cDNAs described here, six plant cDNAs encoding small GTP-binding proteins have been isolated. The very significant amino acid sequence homology may indicate that these proteins are the plant counterplants of the *ypt/rab* type proteins. Complementation experiments with the described plant genes can support this hypothesis and provide further information about the biological function of these plant genes. At present, by analogy to the *ypt/rab* proteins we anticipate similar biological roles for the plant GTP-binding proteins. Consequently, it is conceivable that the *Nt-rab5* and *Rha-1* proteins are involved in the endocytic pathway in plant cells. Although endocytosis is poorly understood in plants, it has been demonstrated to occur in root hairs [7]. Likewise, receptor-mediated endocytosis has been described in plant suspension cultures [18]. The high-level expression of the *A. thaliana Rha-1* gene in roots correlates well with these findings and with the proposed function of this protein in endocytosis [2]. Interestingly, the *Nt-rab5* gene, encoding a protein which displays 78% indentity to the *Rha-1* protein, exhibits a different expression pattern. The tobacco gene shows the highest level of expression in petals, stigmas and anthers but it is also expressed at a fairly high level in root and stem. Despite the different expression patterns (which may represent only different growth conditions) we assume that *Nt-rab5* protein is the tobacco equivalent of the *A. thaliana Rha-1* gene product and that probably both proteins are the functional counterparts of the *Rab5* proteins of mammals.

It is more difficult to assess the function of those plant genes, including the *Np-ypt3*, which show the highest homology to the *S. pombe ypt1* protein. In contrast to the relatively well characterized *ypt3* and *sec4* proteins the physiological function of the *S. pombe ypt3* protein is unknown. Complementation experiments with the *S. pombe ypt3* protein has a unique physiological role in yeast cells [24]. The *O. sativa rgp1* gene with 65% homology to the *S. pombe ypt3* gene showed developmentally regulated expression in young, growing *O. sativa* seedlings; moreover, its expression was strongly reduced in 5-azacytidine-induced dwarf plants. Based on the correlation between the dwarf phenotype and the suppressed expression of the *rgp1* gene, it has been speculated that the *rgp1* type proteins are engaged in transporting essential compounds (i.e., lectins, hemicellulose) to the outside of the plasma membrane [31]. Although this hypothesis is attractive we are hesitant to suggest a similar role for the *Np-ypt3* gene. First, the C-terminus region of the *O. sativa* protein is very divergent from that of the *Np-ypt3* protein. In fact, within the last 30 amino acids the homology is reduced to only 30%. The lack of obvious homology in the hypervariable C-terminus region which was shown to act as targeting signal in mammalian systems suggests a different location, i.e. different functions for these proteins. Second, the tissue-specific distribution of the *Np-ypt3* mRNA is very characteristic in flowering,

mature tobacco plants. As we have described previously, its expression is markedly higher in petals, stamens and stigmas than in other tissues, while its expression is low in leaves and in young (14-day-old) tobacco seedlings. There is no information available yet about the tissue specificity of the *O. sativa rgp1* gene expression. Based on the available data, we suggest the following speculative model. The tobacco *Np-ypt3* and *O. sativa rgp1* proteins are probably both engaged in transporting cell wall components but differ due to distinct tissue specificities and developmental expression patterns. We are presently raising antibodies against the *Nt-rab5* and *Np-ypt3* proteins. We hope that experiments using these antibodies will facilitate the localization of these proteins to specific intracellular compartments.

Acknowledgements

We thank D.W. Kirk for computer analysis and excellent technical help. We are also indebted to F. Meins Jr. and to T. Hohn for the critical reading of the manuscript.

References

1. Adari H, Lowy DG, Willumsen BM, Der CJ, McCormick F: Guanosine triphosphatase activating protein (GAP) interacts with the p21 *ras* effector binding domain. Science 240: 518–521 (1988).
2. Anuntalabhochai S, Terryn N, Van Montagu M, Inze D: Molecular characterization of an *Arabidopsis thaliana* cDNA encoding a small GTP-binding protein, *Rha-1*. Plant J 1: 167–174 (1991).
3. Baker D, Wuestehube L, Scheckman R, Bostein D, Segev N: GTP-binding *YPT1* protein and Ca^{2+} function independently in a cell-free protein transport reaction. Proc Natl Acad Sci USA 87: 355–359 (1990).
4. Balch WE: Small GTP-binding proteins in vesicular transport. Trends Biochem Sci 15: 473–477 (1989).
5. Barbacid M: *ras* genes. Annu Rev Biochem 56: 779–827 (1987).
6. Blum W, Hirsch KD, Schultz G, Weiler EW: Identification of GTP-binding proteins in the plasma membrane of higher plants. Biochem Biophys Res Commun 156: 954–959 (1988).
7. Bonnet HT, Newcomb EH: Coated vesicles and other cytoplasmic components of growing root hairs of radish. Protoplasma 62: 59–75 (1966).
8. Chavrier P, Parton RG, Hauri HP, Simons K, Zerial M: Localization of low molecular weight GTP-binding proteins to exocytic and endocytic compartments. Cell 62: 317–329 (1990).
9. Chavrier P, Gorvel JP, Stelzer E, Simons K, Guenberg J, Zeiral M: Hypervariable C-terminal domain of *rab* proteins acts as a targeting signal. Nature 353: 769–772 (1991).
10. Drivas GT, Shih A, Coutavas E, Rush MG, D'Eustachio P: Characterization of four novel *ras*-like genes expressed in a human teratocarcinoma cell line. Mol Cell Biol 10: 1793–1798 (1990).
11. Gallwitz D, Donath C, Sander C: A yeast gene encoding a protein homologous to the human *c-has/bas* proto-oncogene product. Nature 306: 704–707 (1983).
12. Gorvel JP, Chavrier P, Zerial M, Gruenberg J: *rab5* controls early endosome fusion *in vitro*. Cell 64: 915–925 (1991).
13. Goud B, Zahraoui A, Tavitian A, Saraste J: Small GTP-binding protein associated with Golgi cisternae. Nature 34: 553–556 (1990).
14. Gray JC, Kung SD, Wildman SG, Sheen SJ: Origin of *Nicotiana plumbaginifolia* detected by polypeptide composition of Fraction I protein. Nature 252: 226–227 (1974).
15. Hall A: The cellular functions of small GTP binding proteins. Science 249: 635–640 (1990).
16. Hasunuma K, Furugawa K, Funadera K, Kubota M, Watanabe M: Partial characterization and light induced regulation of GTB-binding proteins in *Lemna pauciostata*. Photochem Photobiol 46: 531–535 (1987).
17. Haubruck H, Engelke U, Mertins P, Gallwitz D: Structural and functional analysis of *YPT2*, an essential *ras*-related gene in the fission yeast *Schizosaccharomyces pombe* encoding a SEC4 protein homologue. EMBO J 9: 1957–1962 (1990).
18. Horn MA, Heinstein PF, Low PS: Receptor-mediated endocytosis in plant cells. Plant Cell 1: 1003–1009 (1989).
19. Laemmli UK: Change of structural proteins during assembly of the head of the bacteriophage T4. Nature 227: 680–685 (1970).
20. Lowy DR, Willumsen BM: New clue to *Ras* lipid glue. Nature 341: 384–385 (1989).
21. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1982).
22. Matsui M, Sasamoto S, Kuneida T, Nomura N, Ishizaki R: Cloning of *ara*, a putative *Arabidopsis thaliana* gene homologous to the *ras*-related family. Gene 76: 331–339 (1989).
23. McCormick F: *ras* GTPase activating protein: signal transmitter and signal terminator. Cell 56: 5–8 (1989).
24. Miyake S, Yamamoto M: Identification of *ras*-related, *YPT* family genes in *Schizosaccharomyces pombe*. EMBO J 7: 1417–1422 (1990).

25. Molenaar CMT, Prange R, Gallwitz D: A carboxyl-terminal cysteine residue is required for palmitic acid binding and biological activity of the *ras*-related yeast *YPT1* protein. EMBO J 7: 971–976 (1988).
26. Nagy F, Kay SA, Chua N-H: The analysis of gene expression in transgenic plants. In: Gelvin SB, Schilperoort RA (eds) Plant Gene Research Manual, pp. 1–29. Kluwer Academic Press, Dordrecht, Netherlands (1987).
27. Nagy F, Kay SA, Chua N-H: Gene regulation by phytochrome. Trends Genet 4: 37 (1988).
28. Pai EF, Kabsh W, Krengel U, Holmes KC, John J, Wittinghofer A: Structure of the guanine-nucleotide-binding domain of the *Ha-ras* oncogene product p21 in the triphosphate conformation. Nature 341: 208–214 (1989).
29. Palmè K, Diefenthal T, Sander C, Vingron M, Schell J: Identification of guanine nucleotide binding proteins in plants: structural analysis and evolutionary comparison of the *ras*-related *YPT*-gene family from *Zea mays*. In: Bosch L, Kraal B, Parmeggiai A (eds) The Guanine Nucleotide Binding Protein: Common Structural and Functional Properties, pp. 272–284. Plenum Press, New York (1989).
30. Salminen A, Novick PJ: A *ras*-like protein is required for a post-Golgi event in yeast secretion. Cell 49: 527–538 (1987).
31. Sano H, Youssefian S: A novel *ras*-related *rgp1* gene encoding a GTP-binding protein has reduced expression in 5-azacytidine-induced dwarf rice. Mol Gen Genet 228: 227–232 (1991).
32. Schmitt HD, Wagner P, Pfaff E, Gallwitz D: The *ras*-related *YPT1* gene product in yeast: a GTP-binding protein that might be involved in microtubule organization. Cell 47: 401–412 (1986).
33. Segev N, Mulholland J, Botstein D: The yeast GTP-binding *YPT1* protein and a mammalian counterpart are associated with the secretion machinery. Cell 52: 915–924 (1988).
34. Sigal IS, Gibbs JB, D'Alonzo JS, Temeles GL, Wolanski BS, Socher SH, Scolnick EM: Mutant *ras*-encoded proteins with altered nucleotide binding exert dominant biological effects. Proc Natl Acad Sci USA 83: 952–956 (1986).
35. Smith DB, Johnson SK: Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-*S*-transferase. Proc Natl Acad Sci USA 83: 8703–8707 (1986).
36. Touchot N, Chardin P, Tavitian A: Four additional members of the *ras* gene superfamily isolated by an oligonucleotide strategy: Molecular cloning of YPT related cDNA from a rat brain library. Proc Natl Acad Sci USA 83: 952–956 (1987).
37. Zahraoui A, Touchot N, Chardin P, Tavitian A: The human *Rab* genes encode a family of GTP-binding proteins related to yeast YPT1 and SEC4 products involved in secretion. J Biol Chem 264: 12394–12401 (1989).