Genet. Res., Camb. (1997), **69**, *pp.* 11–15 *With 6 text-figures*

Mutational analysis of the *rolA* gene of *Agrobacterium rhizogenes* in tobacco: function of the *rolA* pre-mRNA intron and *rolA* proteins defective in their biological activity

A. SPENA[1]* AND K. LANGENKEMPER[2]

[1] *Department of Agricultural & Industrial Biotechnology, Faculty of Science, University of Verona, Strada Le Grazie, 37134 Verona, Italy*
[2] *MPI für Züchtungsforschung, Carl-von-Linne' weg 10, Cologne, Germany*

(*Received 21 July 1996 and in revised form 18 September 1996*)

Summary

The *rolA* gene of *Agrobacterium rhizogenes* contains in its untranslated leader region a spliceosomal intron, which is spliced in *Arabidopsis* and in *Nicotiana tabacum*. Expression under the control of the 35S promoter from cauliflower mosaic virus of a *rolA* gene derivative defective in splicing still causes alterations of growth in transgenic tobacco plants. Splicing of *rolA* mRNA is required for efficient expression of the *rolA* phenotype *in vivo*. Moreover, splicing is required for efficient *in vitro* translation of the *rolA* mRNA. In contrast, expression of a *35S-rolA* gene derivative with the ATG initiation codon replaced by ATA does not cause any phenotypical alteration. Mutations leading to amino acid substitutions at positions 37 and 40 of the *rolA* coding region were isolated as null mutants in *Arabidopsis* plants transgenic for the *rolA* gene. However, when expressed in tobacco under the control of the 35S promoter, they cause a *rolA* phenotype reduced in the expressivity of its traits. The molecular characterization of *rolA* mutants might be useful for understanding the biochemical function of the rolA protein.

1. Introduction

The *rolA* gene from the Ri plasmid A4 of *Agrobacterium rhizogenes* is one of the T_L-DNA genes transferred from bacteria to plant cells, and is involved in the pathogenesis of hairy-root disease (White *et al.*, 1985). When expressed by itself, the *rolA* gene causes developmental alterations in transgenic plants (for review see Michael & Spena, 1995).

Although of bacterial origin, the *rolA* gene contains in its untranslated leader region an intron, which is spliced in *Arabidopsis* (Magrelli *et al.* 1994). In *Arabidopsis*, splicing is required for *rolA* expression because mutations at the 3′ splice site result in transgenic plants without any phenotypical alteration, and *in vitro* translation of *rolA* pre-mRNA having a defective 3′ splice site showed no detectable rolA peptide (Magrelli *et al.*, 1994). Since no differences in the steady-state level of *rolA* mRNA were detected in the mutants, the data indicate that the intron has an inhibitory effect on expression, most probably due to inhibition of *in vitro* translation.

* Corresponding author. Tel. +39-45-8098919. Fax: +39-45-8098929.

In *Arabidopsis*, the initiation of transcription has been mapped 100 bp before the ATG initiation codon. To test whether splicing of the *rolA* intron is absolutely necessary for expression, the region $-100/+330$ (spanning the whole untranslated leader region and the whole coding region of the *rolA* gene, with the A of the ATG initiation codon being $+1$) of the *rolA* gene was expressed in transgenic tobacco plants under the control of the cauliflower mosaic virus (CaMV) 35S promoter and *nos* signals for termination of transcription. The choice of the strong CaMV 35S promoter is based on the observation that expression of the *rolA* gene under the control of the 35S promoter causes a higher expressivity of the traits (Michael & Spena, 1995). Consequently, the expression of the mutated derivatives of the *rolA* gene under the control of the 35S promoter would facilitate the detection of residual activity in mutated *rolA* genes. Indeed, analysis of the transgenic plants shows that a *rolA* mutant defective in splicing is curtailed in its expression and yet is still able to cause alterations of growth.

The function of the rolA protein is unknown. Mutants represent useful tools for understanding the function of a gene. Point mutations within the *rolA* coding region, isolated in *Arabidopsis* as null

mutants (Dehio & Schell, 1993), were characterized at the molecular level, and further analysed in transgenic tobacco plants to test whether the mutations had completely abolished or reduced protein function.

2. Materials and methods

(i) *Construction of chimeric genes and transgenic plants*

Standard procedures were used for recombinant DNA work (Maniatis *et al*., 1989). DNA regions spanning nucleotides −100/+330 were amplified by the polymerase chain reaction (PCR) from the *rolA* gene and its mutants (Dehio & Schell, 1993), and subcloned, as *Bam*H I/*Pst* I fragments under the control of the 35S promoter from CaMV and *nos* termination signals, into the plasmid bluescript KS+. The constructs were checked by DNA sequencing, subcloned as *Eco*R I fragments into the binary vector pPCV002 (Koncz & Schell, 1986) and transferred to *Agrobacterium tumefaciens* strain GV3101 (MP90RK) by electroporation.

Tobacco plants transgenic for the different constructs were raised by leaf-disc transformation using standard procedures (Horsch *et al*., 1985). Independent transgenic plants were screened for expression of the *rolA* gene, or its mutant derivatives, by northern blot analysis. The number of kanamycin-resistant plants expressing the chimeric genes was: 6 of 10 *35S-rolA* plants; 5 of 10 *35S-rolA.1* plants; 12 of 15 *35S-rolA.3* plants; 7 of 21 *35S-rolA.6* plants; 7 of 9 *35S-rolA.7* plants; 13 of 29 *35S-rolA.4(1)* plants. Three plants transgenic for the pPCV002 vector were raised as negative controls. F_1 analysis (selfing) was performed with two clones transgenic for the *35S-rolA.7* gene.

(ii) *RNA extraction and northern blot analysis*

Poly(A+) RNA was extracted, with DynaBeads according to the manufacturer's instructions (Dynal). Agarose–formaldehyde gel (1·2%) electrophoresis, northern blot analysis and hybridization were performed using standard protocols (Maniatis *et al*., 1989).

(iii) *RT-PCR analysis*

Poly(A+) RNA was extracted and purified with Dynabeads. Reverse transcriptase (RT; Superscript II, BRL) reactions were primed with oligo-dT (15mer, Boehringer Mannheim) according to the recommendations of the supplier (BRL). After G tailing with terminal transferase (Boehringer Mannheim), the cDNA was used as a template for PCR primed with the oligonucleotides 5′CCCCCCCCCCCCCA-ACGCTTCAA3′ and 5′AAGACCGCCAGCCA-CGTGCXGTATT3′. (A similar analysis was performed with poly(A+) RNA extracted from *Arabidopsis* plants, and in this case 12 independent clones were sequenced. Also in *Arabidopsis* the GT at position −86 is the most common 5′ splice site, although in this species the GTs at position −73 and −76 are also used; data not shown).

(iv) In vitro *translation of mRNA, immunoprecipitation and PAGE*

In vitro translation of mRNA, immunoprecipitation and PAGE were carried out as previously described (Magrelli *et al*., 1994).

3. Results

(i) *Splicing of the* rolA *pre-mRNA in tobacco is required for efficient expression*

To provide data concerning the role played by the *rolA* intron, the region −100/+330 (the A of the ATG initiation codon being +1) of the *rolA* gene was expressed under the control of the CaMV 35S promoter and *nos* signals for termination of transcription (*35S-rolA*; Fig. 1). Plants expressing the *35S-rolA* gene showed developmental alterations typical of *rolA* gene action (Fig. 2*a*). Northern blot analysis of poly(A+) RNA extracted from these plants showed a single band (Fig. 3*a*, lane 2) hybridizing to a *rolA* probe. DNA sequencing of 12 cloned RT-PCR products showed that all products corresponded to *rolA* mRNA spliced by using the AG at position −3 as 3′ splice site, and the GT at position −86 as 5′ splice site (Fig. 4). Thus, in tobacco, *rolA* pre-mRNA splicing removes an intron of 85 bases (from the G at position −86 to the G at position −2), leaving an untranslated leader region of 15 bases.

Mutations at 3′ splice sites cause either accumulation of unspliced pre-mRNAs and/or activate cryptic splice sites located either upstream or downstream of the canonical AG site (Sharp, 1987). Expression of a *35S-rolA* gene derivative (*35S-rolA.3*; Fig. 1) that has a defective 3′ splice site (AG mutated to AA) caused alterations, typical, albeit reduced in expressivity, of *rolA* gene action (Fig. 2*a*). Poly(A+) RNA extracted from *35S-rolA.3* transgenic plants showed transcripts of more than one size (650 and 740

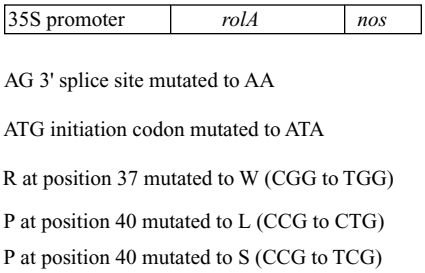

Fig. 1. Schematic drawing of the *35S-rolA* chimeric gene and its mutant derivatives.

(*a*)

(*b*)

Fig. 2. Phenotype of tobacco plants transgenic for the *35S-rolA* chimeric gene and its mutant derivatives. (*a*) From left to right: plant transgenic for the pPCV002 vector alone, for *35S-rolA*, for *35S-rolA.3*, and for *35S-rolA.7* chimeric genes. (*b*) From left to right: plant transgenic for the pPCV002 vector alone, for *35S-rolA.1*, for *35S-rolA.6* and for *35S-rolA.4*(*1*) chimeric genes.

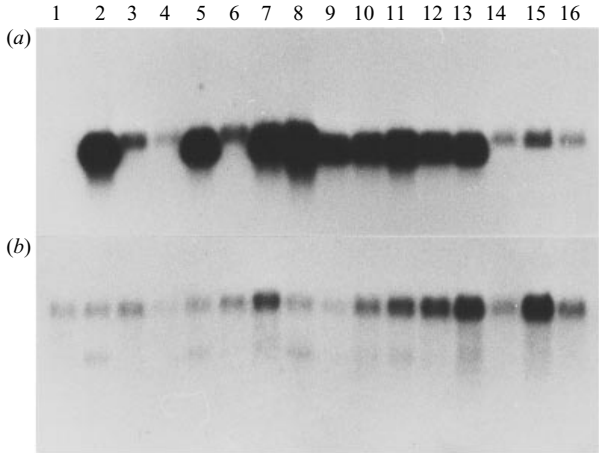

Fig. 3. Northern blot analysis of poly(A+) RNA extracted from transgenic plants. Poly(A) RNA was separated on a 1·2% agarose–formaldehyde gel, transferred to membrane and hybridized with a probe specific for the *rolA* coding region (panel *a*), and with a probe specific for mRNA encoding the S6 ribosomal protein (EMBL Data Bank, acc. number X68050; panel *b*). Lane 1, plant transgenic for the pPCV002 binary vector alone; lane 2, plant transgenic for the *35S-rolA* chimeric gene, lanes 3 and 4, transgenics for the *35S-rolA.1* gene; lanes 5 and 6, transgenics for the *35S-rolA.6* gene; lanes 7 and 8, transgenics for the *35S-rolA.7* gene; lanes 9–16, transgenics for the *35S-rolA.4*(*1*) gene.

bases) when hybridized to a *rolA*-specific probe (Fig. 5, lanes 1–10) and RT-PCR analysis of the same polyA+ RNA identified two different cDNA products: one type (14 clones) corresponded to unspliced pre-mRNA, while the second type (20 clones) corresponded to a *rolA* mRNA spliced using the same GT (at position −86) and, as 3′ splice site, the AG at position +9, instead of the AG at position −3 (Fig. 4). The activation of the cryptic splice site removes an intron (96 bases long) which includes the AUG initiation codon of the *rolA* mRNA. *In vitro* translation of *rolA* poly(A+) RNA extracted from *35-rolA.3* plants was reduced (at least 20–40 times; compare lane 3 with lanes 4 and 6, Fig. 6). Thus, efficient and correct splicing of the *rolA* intron is required for efficient *in vitro* translation of *rolA* mRNA, and for efficient expressivity of the *rolA* traits *in vivo*.

The residual phenotype could be due to translation of unspliced *rolA* pre-mRNA or/and to translation of aberrantly spliced *rolA* mRNA. The *rolA* transcript aberrantly spliced has the potential, by using an internal AUG, to code for a rolA peptide with its first 10 amino acids deleted. To test whether a *rolA* gene coding for a rolA peptide truncated of its first 10 amino acids is able to cause growth alterations in transgenic plants, a *35S-rolA* derivative with the ATG initiation codon mutated to ATA (*35S-rolA.7*; Fig. 1) was expressed in transgenic plants. Despite their high steady-state level of *rolA* mRNA (Fig. 3*a*, lanes 7 and 8), tobacco plants transgenic for the *35S-rolA.7* gene did not show any alteration of growth (plant *35S-rolA.7*; Fig. 2*a*). *rolA* mRNA extracted from these plants was properly spliced (Fig. 4), while *in vitro* translation and immunoprecipitation of mRNA extracted from plants transgenic for the *35S-rolA.7* gene showed a barely detectable level of rolA peptide (Fig. 6, lane 5). Thus, the use of the internal

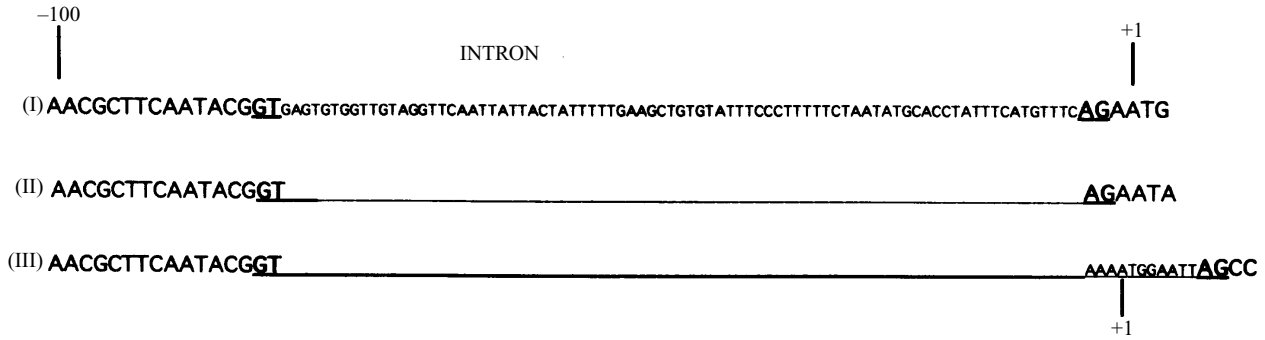

Fig. 4. DNA sequence of spliced *rolA* mRNA. The data were obtained by sequencing RT-PCR products of *rolA* poly(A+) RNA from the following type of transgenic plants; (I) *35S-rolA*, (II) *35S-rolA.7*; (III) *35S-rolA.3*.

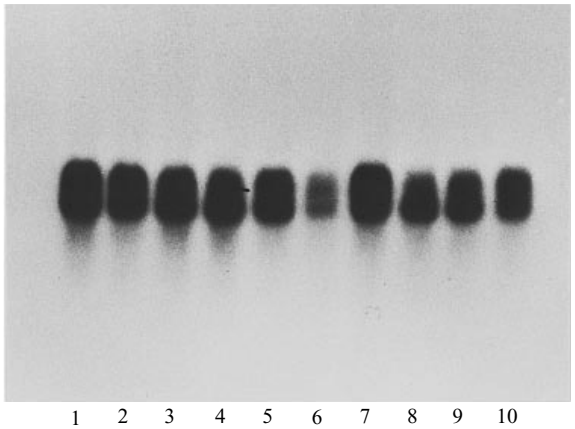

Fig. 5. Northern blot analysis of poly(A+) RNA extracted from 10 independent plant clones transgenic for the *35S-rolA.3* chimeric gene. From lane 1 to lane 10: clone e, f, g, i, k, l, 1, 2, 3 and 4. The northern blot was hybridized to a probe spanning the *rolA* coding region (300 bp). Approximately the same amount of poly(A+) RNA was loaded in each lane, with the exception of lane 6 which contains a lower amount (approximately 1/4).

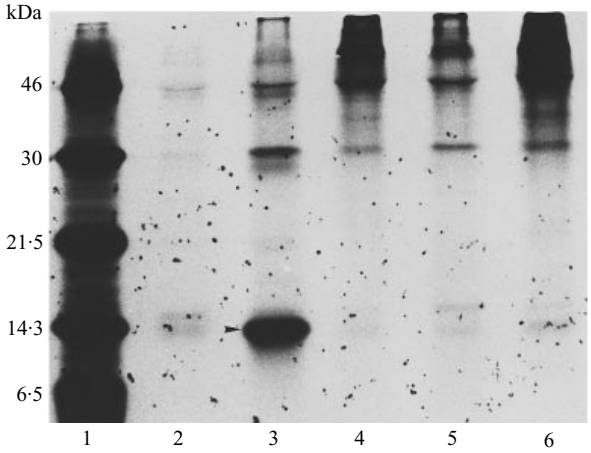

Fig. 6. Immunoprecipitation with affinity-purified anti-rolA antibody and protein A–sepharose of products from an *in vitro* translation in wheat germ system of poly(A+) RNA extracted from tobacco plants. Lane 2, wild-type; lane 3, transgenic for the *35S-rolA* gene; lanes 4 and 6, transgenic for the *35S-rolA.3* gene; lane 5, transgenic for the *35S-rolA.7* gene. Lane 1, Rainbow ^{14}C methylated protein marker (Amersham). The following amounts of protein were immunoprecipitated: wild-type, 1·3 million c.p.m.; *35S-rolA*, 1·3 million c.p.m.; *35S-rolA.3,* 1·3 million c.p.m. in lane 4 and 3·9 million c.p.m. in lane 6; *35S-rolA.7*, 0·8 million c.p.m. The gel was treated with Enhance (Amersham) and exposed for 5 days. The arrow indicates the *rolA* product.

initiation codon curtails *rolA* mRNA translation and abolishes *rolA* function.

This result favours the interpretation that the residual phenotype observed in tobacco plants defective for *rolA* pre-mRNA splicing is due to residual translation of *rolA* pre-mRNA. This interpretation is also consistent with the knowledge that mutations affecting splicing increase the proportion of pre-mRNAs translated *in vivo* (Legrain & Rosbach, 1989).

(ii) *Characterization of amino acid substitutions affecting* rolA *function*

Mutational analysis of the *rolA* gene in *Arabidopsis* (Dehio & Schell, 1993; Magrelli *et al.*, 1994) has shown that two amino acids (R^{37} and P^{40}), defined by three mutations replacing R^{37} with W and P^{40} with either S or L, are relevant for *rolA* protein function. When expressed in *Arabidopsis* under the control of its promoter, the three mutations cause a null phenotype (i.e. transgenic plants do not show any phenotypical alteration, although the steady-state level of *rolA* mRNA is unaffected (Dehio & Schell, 1993). Tobacco plants expressing *35S-rolA* gene derivatives (Fig. 1) having either the R^{37} replaced with W (*35S-rolA.4(1)* gene; Fig. 1*b*), or the P^{40} replaced either with S (*35S-rolA.1*; Fig. 1*b*) or with L (*35S-rolA.6* gene; Fig. 1*b*) displayed developmental alterations typical of *rolA* gene action. In independent transgenic plants, the expressivity of the traits varied with the level of expression of *rolA* mRNA (Fig. 3*a*); however, it was always weaker than that displayed by *35S-rolA* plants. Thus, the three amino acid substitutions do not abolish, but reduce *rolA* function.

4. Discussion

Bacteria belonging to the genus *Agrobacterium* represent an example of horizontal transfer of genetic information from prokaryotes to eukaryotes (for review, see Winans, 1992). One of the genes transferred

from *Agrobacterium* to the plant cell is the *rolA* gene (for review see Michael & Spena, 1995), which is the only gene of bacterial origin containing a spliceosomal intron (Magrelli *et al*., 1994). The intron is located in the untranslated leader region of *rolA* pre-mRNA and is spliced in both *Arabidopsis* and tobacco. The intron must be removed for proper expression in both plants. Its presence dramatically inhibits, but does not abolish, *in vivo* expression of the *rolA* phenotype and *in vitro* translation of *rolA* mRNA.

Mutants of the *rolA* gene in which the ATG initiation codon is changed to ATA are not able to cause any alterations of growth in transgenic plants, even when expressed under the control of the 35S promoter from cauliflower mosaic virus. Such mutation would alter protein translation by moving the initiation of translation to the next AUG located after 30 nt. Thus, the null phenotype most probably results from the use of a less efficient initiation codon and inactivity of the truncated protein (i.e. 10 amino acids shorter at the amino-terminal). The other three mutations isolated as null mutants in *Arabidopsis* lead to substitution of R^{37} with W and P^{40} with either S or L. Although isolated as null mutants in *Arabidopsis*, when expressed in tobacco under the control of the 35S promoter, the mutated *rolA* gene derivatives are still able to cause developmental alterations. The low level of expressivity of the traits, despite the high steady-state level of *rolA* mRNA, indicates that the amino acid substitutions do not abolish, but dramatically reduce, the activity of the rolA protein.

These two amino acids fall in a 41 amino acid domain which is highly conserved (85% identity) between rolA proteins (58% identity overall) from two different *A. rhizogenes* strains (i.e. A4 and 8196; Michael & Spena, 1995). Furthermore, these two amino acids fall in a stretch of 27 amino acids which has 8 amino acids identical and 5 conserved compared with the indoleacetic acid–lysine synthetase (*iaaL*) from *Pseudomonas syringae* pv. *savastanoi* (Michael & Spena, 1995). The phenotype caused in tobacco plants by the expression of the *iaaL* gene under the control of the 35S promoter is similar to that shown by tobacco plants transgenic for the *rolA* gene (Spena, 1993).

We thank Elke Kemper for assistance in raising the transgenic tobacco plants.

References

Dehio, C. & Schell, J. (1993). Stable expression of a single-copy *rolA* gene in transgenic *Arabidopsis thaliana* plants allows an exhaustive mutagenic analysis of the transgene-associated phenotype. *Molecular and General Genetics* **241**, 359–366.

Horsch, R., Fraley, R., Rogers, S., Sanders, P., Lloyd, A. & Hoffmann, W. (1984). Inheritance of functional foreign genes in plants. *Science* **223**, 496–499.

Koncz, C. & Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular and General Genetics* **204**, 383–396.

Legrain, P. & Rossbach, M. (1989). Some *cis*- and *trans*-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* **57**, 573–583.

Magrelli, A., Langenkemper, K., Dehio, C., Schell, J. & Spena, A. (1994). Splicing of the *rolA* transcript of *Agrobacterium rhizogenes* in *Arabidopsis*. *Science* **266**, 1988–2000.

Maniatis, T., Frisch, E. F. & Sambrook, J. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Michael, T. & Spena, A. (1995). The plant oncogenes *rolA*, *B* and *C* from *Agrobacterium rhizogenes*: effects on morphology, development and hormone metabolism. In *Agrobacterium Protocols*, *Methods in Molecular Biology*. Clifton, NJ: Humana Press.

Sharp, P. A. (1987). Splicing of messenger RNA precursors. *Science* **235**, 766–771.

Spena, A. (1993). Transgenic plants altered in phytohormone metabolism. *Acta Botanica Gallica* **140**, 693–700.

White, F. F., Taylor, B. H., Huffman, G. A., Gordon, M. P. & Nester, E. W. (1985). Molecular and genetic analysis of the transferred DNA regions of the root inducing plasmid of *Agrobacterium rhizogenes*. *Journal of Bacteriology* **164**, 33–44.

Winans, S. (1992). Two-way chemical signaling in *Agrobacterium*–plant interactions. *Microbiological Reviews* **56**, 12–31.