CHAPTER 10

Epitope Tagging for the Detection of Fusion Protein Expression in Transgenic Plants

Mark J. Guiltinan and Lauren McHenry

Department of Horticulture and The Biotechnology Institute
The Center for Gene Expression
Pennsylvania State University
University Park, Pennsylvania 16802

I. Introduction

Methods for the genetic transformation of plants have made possible new experimental approaches, revolutionizing the science of plant biology. This technology is also the basis for the production of genetically engineered plants with enhanced or novel genetic traits (Gasser and Fraley, 1992). When conducting genetic transformation studies, it is necessary to have available molecular probes to enable the detection of the transgene and its RNA and protein products (trans-message and trans-protein) via nucleic acid hybridization, antibody local-

METHODS IN CELL BIOLOGY, VOL. 49

ization and/or various functional assays. In some cases, it may also be necessary to distinguish the trans-molecules from any similar cross-reactive endogenous factors.

When existing molecular probes are not available or when endogenous factors interfere with detection, it may be advantageous to engineer into a transgene a short DNA sequence specifying a known polypeptide epitope (an epitope tag). A monoclonal antibody with strong binding specificity for the epitope tag can then be used to detect sensitively the expression of a trans-protein using standard techniques (Western blot analysis, ELISA, or *in situ* immunolocalization). Although larger molecular fusions have also been used for detection, one advantage of an epitope tag is its small size, minimizing the potential for disruption of the proper folding and/or activity of the trans-protein.

One well-characterized epitope is a 10-amino acid sequence derived from the human c-*myc* gene (containing sequences specifying amino acids 410–419). This polypeptide is specifically recognized by the monoclonal antibody Mycl-9E10 (Evan *et al.,* 1985). This epitope has been used to detect the expression of fusion protein in transformed animal and yeast cells. (Ellison and Hochstrasser, 1991; Fowlkes *et al.,* 1992; Munro *et al.,* 1986; Peculis and Gall, 1992). A similar approach utilizing an 8-amino acid epitope tag derived from the human tenascin gene has been used for detecting recombinant RNA binding protein in transfected protoplasts of *Nicotiana plumbaginifolia* (Mieszczak *et al.,* 1992). We describe here the development of the c-*myc* epitope tag system for the detection of trans-protein expression in transgenic tobacco plants. An expression vector for plant c-*myc* epitope tagging was constructed and introduced into tobacco plants via *Agrobacterium*-mediated transformation. Our results show that c-*myc* epitope tagged trans-protein can be detected in extracts from as little as 5 mg of leaf tissue. The Mycl-9E10 monoclonal antibody shows no cross-reactivity to endogenous tobacco or *Arabidopsis thaliana* proteins. Furthermore, the antibody can be used to localize expression in various tissues using the tissue print method (Cassab and Varner, 1987).

II. Methods

A. c-*myc* Epitope Tag Vector Construction

A plant expression vector was designed for the fusion of the c-*myc* epitope tag to any coding sequence, and for expression of the tagged fusion protein in transgenic plants. The vector was derived from pMON881, with the incorporation of an ATG start codon in frame with DNA sequences specifying 10 amino acids of the human c-*myc* protein flanked by the cauliflower mosaic virus 35S promoter and the nopaline synthase 3′ untranslated region. pLM1 (Fig. 1) carries a fusion gene consisting of the c-*myc* epitope tag and the DNA binding domain (λGC19) of the wheat EmBP-1 gene (Guiltinan *et al.,* 1990). Additional experimental results using the c-*myc*/GC19 transgenic plants will be presented elsewhere.

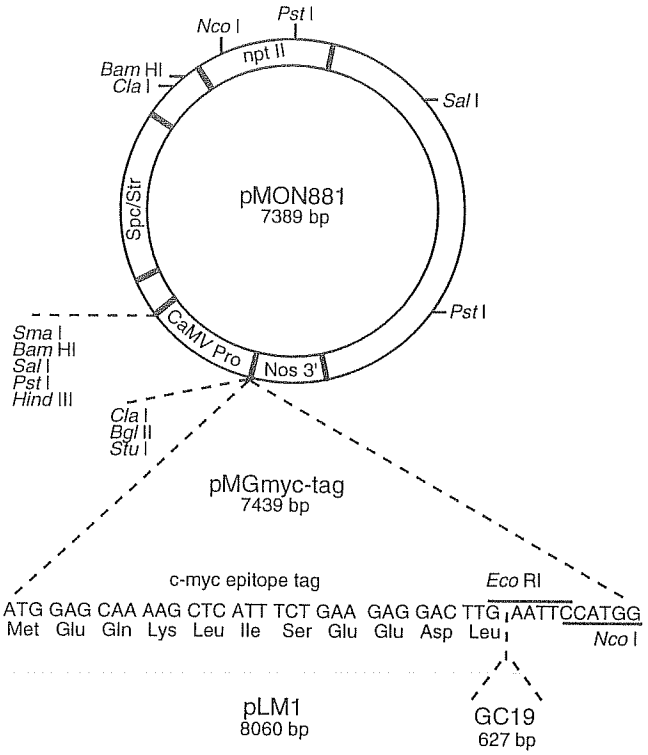

Fig. 1 Plant transformation vector and sequence of c-*myc* epitope tag. Sequences encoding a 30-bp portion of the human c-*myc* gene were inserted into the plant transformation vector, pMON881 (Monsanto; St. Louis, MO), to create pMGmyc-tag. An ATG start codon is present just 5′ to sequences encoding the epitope tag and there is a unique *Eco*RI cloning site directly following the epitope tag. The plant cell selectable marker gene (kanamycin resistance, npt II) is indicated. The DNA binding and dimerization domain (GC19) of the EmBP-1 protein (Guiltinan *et al.,* 1990) was cloned into the *Eco*RI site of pMGmyc tag to produce pLM1 used in the production of transgenic tobacco plants.

Standard protocols were used for DNA isolation and manipulation (Sambrook *et al.,* 1989), and for transformation of *Nicotiana tabacum,* SR1 (Horsch *et al.,* 1988). The *Hind*III/*Spe*I fragment of pMG99.23, containing sequences specifying amino acids 410–419 of the human c-*myc* gene fused to EmBP-1 sequences from the *Eco*RI insert of λGC19 (Guiltinan *et al.,* 1990), was rendered blunt with T4 polymerase and subcloned into the similarly blunt-ended *Eco*RI site of the plant expression vector pMON881 (Monsanto, St. Louis, MO), to yield pLM1 (Fig. 1).

pMG99.23 was constructed by inserting the Klenow filled *Eco*RI insert of λGC19 (Guiltinan *et al.,* 1990) into the *Bam*HI cut and filled vector pAU1 to form pAU2. pAU1 was derived from pBSKSc-*myc* (Peculis and Gall, 1992),

which contains a 30-bp synthetic oligonucleotide, representing a portion of the human c-*myc* coding region (Fig. 1), inserted into the *Hin*dIII/*Eco*RI site of pBluescript KS+ (Promega, Madison, WI). Stop codons in all three reading frames were introduced downstream of the unique *Spe*I site of pBSKSc-*myc* to form pAU1 by cutting, filling, and religating at the *Xba*I site. This change and other critical features of the construct were verified by DNA sequencing. The *Sac*I/*Xho*I fragment of pAU2, containing the short c-*myc* fragment fused to the 5′ end of the GC19 sequences, was subcloned into *Sac*I/*Xho*I-digested pRT101 (Topfer *et al.,* 1987) to create pMG99.23.

B. Plant Transformation

pLM1 was introduced into the *Agrobacterium tumefaciens* strain ABI via triparental mating (Van Haute *et al.,* 1983) utilizing pRK2013 as helper plasmid. Tobacco (*N. tabacum* Petite Havana cv. SR1) leaf discs were transformed with the resulting *A. tumefaciens* strain, and regenerated shoots (R_0) were selected for their ability to root on MSO media (MS salts, Gamborg's vitamins, 1% sucrose, 0.3% phytagel) containing 100 μg/ml kanamycin monosulfate. Stable transformation of nine individual lines was confirmed by segregational analysis of the *npt*II selectable marker in seed lots (F1) from self-fertilized primary transformants plated on germination media (MS salts, Gamborg's vitamins, 0.7% glucose, 0.3% sucrose, 0.3% phytagel) containing 100 μg/ml kanamycin.

C. Protein Extraction, Electrophoresis, and Blotting

Primary transformants (R_0) were self-fertilized and the resulting kanamycin-resistant offspring (F1) were tested for expression of the c-*myc*/GC19 fusion protein by extracting total SDS-soluble protein from 3-week-old seedlings and subjecting it to Western blot analysis (Fig. 2). After 3 weeks of growth on germination media supplemented with 100 μg/ml kanamycin, 10 kanamycin-resistant seedlings of each transgenic line were randomly selected and pooled. Each pool was homogenized with a mini-mortar in a 1.5-ml microcentrifuge tube with a small volume of 2× Laemmli buffer (Laemmli, 1970). The samples were placed in a boiling water bath for 10 min and then microfuged for 10 min at 4°C to remove debris. The volume of each sample was adjusted to a final concentration of 1 mg fresh weight per microliter with the addition of 10× loading dye (0.1% bromphenol blue in 80% glycerol) and water. Protein from roots and leaves of mature soil-grown *Arabidopsis* was prepared in the same manner.

Approximately equal amounts of each sample, including a wild-type SR1 control protein extract, were electrophoresed through a 10% SDS–polyacrylamide gel using a discontinuous buffer system. Separated proteins were electroblotted onto an Immobilon-P transfer membrane (Millipore Corp., Bedford, MA) using a submerged mini trans-blot electrophoretic transfer cell (BioRad; Richmond, CA) according to the manufacturer's instructions by applying 100 V for

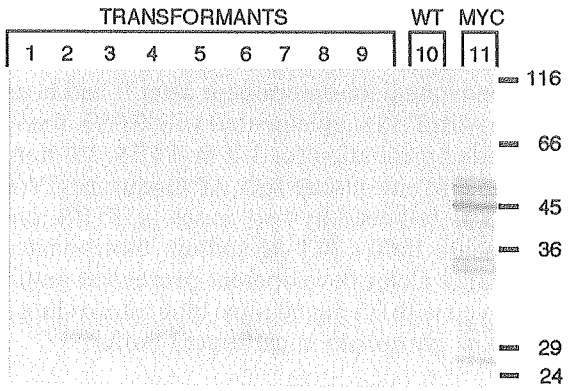

Fig. 2 Western blot analysis of protein extracted from c-*myc*/GC19 transgenic tobacco plants reacted with an antibody directed against the c-*myc* epitope. Total SDS-soluble protein from pools of kanamycin-resistant seedlings from self-crossed primary transformants were separated by polyacrylamide gel electrophoresis and transferred to Immobilon-P. The filter was reacted with a monoclonal antibody (Mycl-9E10) directed against the human c-*myc* protein. Each lane contains an extract equivalent to approximately 5 to 10 mg of fresh weight-seedling tissue. Lanes 1–9, nine independent transformants; Lane 10, wild-type SRI tobacco; Lane 11, c-*myc*-tagged protein derived from expression of the *Ste7* gene of *Saccharomyces cerevisiae* in *Escherichia coli.*

1 h in a transfer buffer composed of 25 m*M* Tris [hydroxymethyl]aminomethane, 192 m*M* glycine, 20% (v/v) methanol.

D. Tissue Printing and Protein Staining

Transgenic F2 seedlings propagated on MSr plates containing 100 μg/ml kanamycin, and control SR1 seedlings on MSr plates without kanamycin, were rinsed with sterile distilled water and blotted onto nitrocellulose (0.45 μm S&S NC, Schleicher & Schuell, Keene, NH) as previously described (Cassab and Varner, 1987). Blots were then stained to confirm protein transfer by a 1-min soak in 0.5% Ponceau-S in 1% acetic acid and destained by several subsequent rinses in distilled water. Immunodetection was performed as described below. After color development, nitrocellulose blots were rinsed in full strength Clorox bleach for 2 min to reduce residual chlorophyll.

E. Antibody Incubations and Immunodetection

The membranes with affixed proteins were equilibrated in TBS (Tris-buffered saline: 20 m*M* Tris, pH 7.5, and 500 m*M* sodium chloride) and then blocked with 5% Carnation instant nonfat dry milk in TBS for 1 h at room temperature. The

primary antibody, Mycl-9E10, was a monoclonal antibody raised against the human c-*myc* protein in mouse cells (Evan *et al.,* 1985) that were grown to confluency in Dulbecco's medium (GIBCO). The membrane was saturated with antibody medium supernatant after it had been blocked as described and rinsed twice with TBS supplemented with 0.05% Tween-20 (TTBS). Two rinses in TTBS preceded incubation for 1 h in TTBS, 5% nonfat dry milk, and 1:7500 dilution of 1 mg/ml anti-mouse IgG AP conjugate (Promega, Madison, WI). The incubation was followed by two rinses in TTBS, one rinse in TBS, and one rinse in carbonate buffer (0.1 *M* sodium bicarbonate, pH 9.8, and 1 m*M* magnesium chloride). Color development proceeded until full color of reacting proteins was achieved with 0.3 mg/ml nitro blue tetrazolium and 0.15 mg/ml 5-bromo-4-chloro-3-indoyl phosphate in carbonate buffer.

III. Results and Discussion

A fusion gene was constructed consisting of the c-*myc* epitope tag and a portion of the wheat EmBP-1 gene under the control of the CaMV 35S promoter (Fig. 1). Transgenic tobacco lines with the tagged gene were generated via Agrobacterium-mediated gene transfer. Protein extracts from nine transgenic and one normal plant were analyzed for the presence of the fusion protein by Western blot analysis (Fig. 2). An antibody directed against the c-*myc* epitope (Mycl-9E10) allowed detection of the c-*myc*/GC19 protein with the expected apparent molecular mass of 30 kDa in crude extracts from four of nine transformants (Fig. 2, lanes 1, 2, 4, and 6). Expression level variation among independent transgenic lines is the most likely explanation for why no fusion protein was detected in five of the plants. No c-*myc* reactive protein was detected, even after prolonged color development, in SR1 wild-type tobacco plants (Fig. 2, lane 10), or in protein extracts from roots or leaves of *Arabidopsis thaliana* (data not shown). Although we have not yet determined the sensitivity of this detection system, we were capable of detecting protein from as little as 5 mg of fresh weight leaf tissue.

Tissue printing allows the simple detection of the amounts of protein or mRNA present at the whole organism level. To test the suitability of the c-*myc*-tag system to tissue printing, whole seedling prints were prepared and reacted with the c-*myc* antibody. The antibody strongly detected expression of the fusion protein in the leaves, roots, and cotyledons of transgenic F1 seedlings (Fig. 3), consistent with the known expression pattern of the CaMV35S promoter. A stronger signal was seen on the periphery of the tissue, perhaps due to more efficient extraction of protein around the edges. No cross-reactivity of the antibody to tissue prints of control, nontransformed tobacco plants was detectable.

IV. Conclusions and Perspectives

The ability of the Mycl-9E10 antibody to detect the 10-amino acid c-*myc* epitope on Westerns blots and tissue prints and the absence of antibody cross-

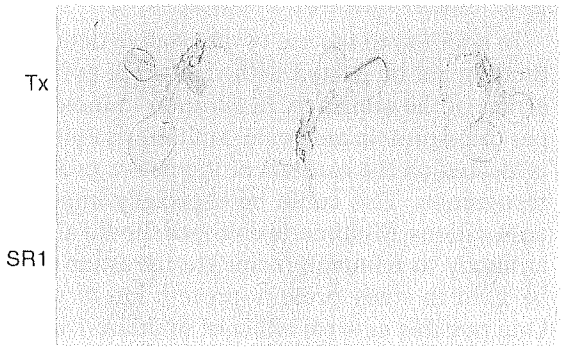

Fig. 3 Tissue prints of c-myc/GC19-transformed tobacco plants reacted with an antibody directed against the c-*myc* epitope. Whole seedlings of either c-*myc*-tagged GC19-transformed (Tx) or wild-type (SR1) lines were blotted onto nitrocellulose and reacted with a monoclonal antibody (Mycl-9E10) directed against the human c-*myc* protein.

reactivity to endogenous tobacco and *Arabidopsis* proteins make this epitope–antibody pair an excellent choice as a general marker for protein expression in transgenic plants. In cases where no enzymatic activity has been demonstrated for a protein, or in which no antibody specific to the encoded protein is availabile, epitope-tagging provides a simple alternative to monitoring protein expression of introduced genes.

Tissue prints of excised leaves can also be used to detect rapidly the expression of the c-*myc* tagged trans-protein in a large population of segregating progeny (data not shown). This is particularly advantageous in the cosegregational analysis of a trans-protein and its presumptive phenotypic effects. Leaf disks taken from transgenic seedlings can be arrayed in a grid and rapidly tested for trans-protein expression. This allows for hundreds of samples to be processed in a single day, thus enabling the screening of sufficient numbers of progeny to confidently test for cosegregation with a phenotypic character scored separately for each seedling.

Epitope tagging is not without its potential pitfalls. The most important consideration is the theoretical possibility that the epitope tag might in some way destabilize a fusion protein, alter or eliminate its biological function, or change its subcellular localization pattern. Functional inactivation is frequently observed in the production of bacterial fusion proteins when large affinity purification tags are used such as β-galactosidase, maltose binding protein, or glutathione S-transferase. In cases where an N-terminal fusion is a problem, it would be necessary to attempt C-terminal or internal fusions, which would be difficult with our present vector. The effect of an epitope on the behavior of a fusion protein is likely to be highly variable and dependent on the structure of the protein under study. It is not currently possible to predict accurately protein structure from sequence data, and thus such effects must be tested on a case by case basis.

A. Potential Improvements and Applications of Epitope Tags

In addition to the use of alternative detection systems, improvements in sensitivity might be gained with the use of epitope multimers, four or more tandem arrays of the sequence. In addition to increasing the number of epitope moieties per trans-protein molecule, multimeric epitopes may help ensure that the epitope is located on the exterior of the native protein and thus is available for antibody recognition. This could be especially important for *in situ* immunolocalization applications. Multimeric epitopes should also improve the ability of a monoclonal antibody to immunoprecipitate the trans-protein. This powerful technique can be used to study protein–protein interactions via coprecipitation experiments. One possible adverse effect of multimerization is that longer fusions have inherently higher probability of causing structural perturbation to the protein and thus increased potential to reduce or eliminate its functional activity.

An additional potential of the epitope tag system is the possibility of using the DNA sequence encoding the c-*myc*-tag as a hybridization probe to detect the transgene or its message in DNA and RNA isolated from transgenic plants. This would be advantageous when endogenous genes or mRNAs interfere with the specific detection of a transgene and/or its message by cross-hybridization to the engineered gene probe. Nucleic acid probes shorter than 30 nucleotides have been successfully used in similar hybridization experiments, so there is no a priori reason this should not be possible. It may also be possible to use such probes to localize the trans-message at a cellular level using *in situ* hybridization. However, the small size of such a probe may reduce the detection sensitivity below acceptable levels or cause high background hybridization. These limitations might be avoided by the use of short primers to amplify the tagged messages *in situ* using PCR. If possible, this would allow for the sensitive detection of the tissue, cellular, and subcellular distribution patterns of trans-gene expression levels at the level of mRNA.

The c-*myc* epitope-tag system provides a sensitive and specific tool for the analysis of gene expression in transgenic plants. The relative ease by which such constructions can be generated, the availability and high specificity of the Mycl-9E10 monoclonal antibody, and the various well-developed immunodetection techniques available make epitope tagging a powerful tool in the study of plant molecular biology.

B. Vector and Antibody Availability

The plant expression vector pMGmyc-tag (Fig. 1), containing the c-*myc* epitope tag, is available from the authors after first obtaining permission for use of the parental plasmid pMON881 from the Monsanto Co. (Dr. Harry Klee, 700 Chesterfield Village Parkway, St. Louis, MO 63198). The c-*myc* antibody cell line (Mycl-9E10.2; ATCC CRL 1729) is available from the American Type Culture Collection c/o Sales and Marketing Department, 12301 Parklawn Drive, Rockville, MD 20852, 1-800-638-6597.

Acknowledgments

We thank Dr. Beverly Erredy (University of North Carolina at Chapel Hill) for supplying c-*myc* tagged *Ste*7 protein, Dr. Richard Cyr and Debra Fisher (Penn State University) for producing the Mycl-9E10 antibody supernatant, and Dr. Harry Klee and the Monsanto Co. for providing the pMON881 vector. This work was supported in part by grants to M.J.G. from the NSF (MCB 920 6095) and from Penn State University College of Agricultural Sciences (CRIS 3278).

References

Cassab, G. I., and Varner, J. E. (1987). Immunocytolocalization of extensin in developing soybean seed coats by immunogold-silver staining and by tissue printing on nitrocellulose paper. *J. Cell Biol.* **105,** 2581–2588.

Ellison, M. J., and Hochstrasser, M. (1991). Epitope-tagged ubiquitin. *J. Biol. Chem.* **266,** 21150–21157.

Evan, G. I., Lewis, G. K., Ramsay, G., and Bishop, M. J. (1985). Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell Biol.* **5,** 3610–3616.

Fowlkes, D. M., Adams, M. D., Fowler, V. A., and Kay, B. K. (1992). Multi-purpose vectors for peptide expression on the surface of the M13 viral surface. *Biotechniques* **13,** 422–428.

Gasser, C. S., and Fraley, R. T. (1992). Transgenic plants. *Sci. Am.* June, 62–69.

Guiltinan, M. J., Marcotte, W. R., and Quatrano, R. S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* **250,** 267–271.

Horsch, R. B., Fry, J., Hoffmann, N., Neidermeyer, J., Rogers, S. G., and Fraley, R. T. (1988). Leaf disc transformation. *In* "Plant Molecular Biology Manual" (S. B. Gelvin, R. A. Schilperoort, and D. P. S. Verma, eds), pp. A5/1–A5/9. Dordrecht, The Netherlands: Kluwer.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

Mieszczak, M., Klahre, U., Levy, J. H., Goodall, G. J., and Filipowicz, W. (1992). Multiple plant RNA binding proteins identified by PCR: Expression of cDNAs encoding RNA binding proteins targeted to chloroplasts in Nicotiana plumbaginifolia. *Mol. Gen. Genet.* **234,** 390–400.

Munro, S., and Pelham, H. R. B. (1986). An Hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* **46,** 291–300.

Peculis, A. B., and Gall, J. G. (1992). Localization of the nucleolar protein NO38 in amphibian oocytes. *J. Cell Biol.* **116,** 1–14.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Topfer, R., Matzeit, V., Gronenborn, B., Schell, J., and Steinbiss, H. H. (1987). A set of plant expression vectors for transcriptional and translational fusions. *Nucl. Acids Res.* **15,** 5890.

Van Haute, E., Joos, H., Maes, M., Warren, G., Van Montagu, M., and Schell, J. (1983). Intergeneric transfer and exchange recombination of restriction fragment cloned in pBR322: A novel strategy for reversed genetics of the Ti plasmids of Agrobacterium tumefaciens. *EMBO J.* **2,** 411–417.