In vitro characterization of a cassette to accumulate multiple proteins through synthesis of a self-processing polypeptide

Jose F. Marcos and Roger N. Beachy*
Department of Cell Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, USA (*author for correspondence)

Received 5 August 1993; accepted in revised form 23 November 1993

Key words: expression cassette, polyprotein processing, tobacco etch virus (TEV), viral proteinase

Abstract

The strategy for processing the polyprotein encoded by plant potyviruses has been mimicked by constructing an expression cassette based on the nuclear inclusion (Nla) proteinase from tobacco etch virus (TEV). This cassette (pPR01), includes the TEV Nla coding region flanked on each side by its heptapeptide cleavage sequence and cloning sites for the in frame insertion of two different open reading frames. pPR01 allows the synthesis, under the control of a single transcriptional promoter, of two proteins in equimolar amounts as part of a polyprotein which is cleaved into individual mature products by the TEV protease. In in vitro reactions the cassette functioned as expected when several different protein-coding sequences were used. The potential uses of pPR01 are discussed.

Introduction

In recent years, development of plant transformation techniques and strategies for enhancing and controlling gene expression have broadened the practical applications of plant biotechnology. However, the potential of these techniques must deal with the problems encountered when more than one transgene is expressed in planta.

Current approaches to express more than one gene in transgenic plants require the use of multiple promoters, which in itself presents problems related to levels of expression from each promoter [16]. For example, the relative levels of expression in potato plants of two genes encoding two viral coat proteins (CP), which were introduced via a single Ti-derived transformation vector, were different in different plant lines [15]. An alternative approach, i.e., re-transforming plants with a second gene, may induce gene-silencing effects [11, 17]. Sexual crossing of different transgenic lines may enhance or inhibit gene expression depending on gene copy number and the nature of gene insertion [13]. Therefore, the relative levels of gene expression can not be predicted but rather are a consequence of experimental variability.

Our laboratory is interested in alternative mechanisms to express multiple genes in a single transgenic line as a way to improve pathogen derived protection against plant viruses. In this context, we have developed a system which allow equimolar accumulation of several proteins under the control of a single transcriptional promoter, thereby avoiding the problems outlined above.

Several plant and animal viruses encode pro-
teinases that cleave the viral polypeptides yielding mature proteins. Plant potyvirus genomes are expressed through the translation of a single polypeptide which is processed to release individual viral proteins [20]. Three viral proteinase activities have been implicated in this processing [7, 24]. One of these, corresponding to the nuclear inclusion (Nla) protein, has been widely studied in the case of tobacco etch potyvirus (TEV) [2, 4], and is responsible for several processing events of the large viral polypeptide. Nla from TEV exerts this function through the recognition and cleavage of a specific heptapeptide [Glu-X-X-Tyr-X-Gln/Gly (or Ser)] [5, 8]. Nla releases itself from the polypeptide in an autoproteolytic reaction [3], and is active both in cis (processing polypeptides in which it is included) and in trans (cleaving in a different polypeptide). The cis protease activity has been assayed with different TEV polypeptides produced in vitro which contained Nla and either naturally occurring or mutated versions of the cleavage sequence [2, 4]. Protease activity in trans has been studied using as substrates TEV polypeptides that were labeled in vitro and incubated with Nla either extracted from infected plants [4, 5, 8, 9] or synthesized in in vitro translation reactions [2, 4, 18].

We have taken advantage of this well characterized proteinase activity to develop an expression cassette based on the TEV-Nla protein. This cassette allows the synthesis of two proteins in equimolar amounts as part of a polypeptide that is cleaved into individual mature proteins by the Nla proteolytic activity. The in vitro characterization of this expression system demonstrated that it functions as expected when different proteins are used.

Materials and methods

Cloning strategies

Recombinant DNA manipulation and Escherichia coli formation were carried out according to existing protocols [22]. The DNA inserts used for the assembly of the different constructs were obtained by the polymerase chain reaction (PCR, Perkin Elmer Cetus). The expression cassette pPRO1 (see Results and Fig. 1) was assembled in pBluescript II KS (+) (Stratagene) under the transcriptional control of a T7 promoter. Nla and 5’-non-translated (5-NTR) sequences from TEV were obtained by PCR using as DNA template a full-length TEV cDNA clone (kindly provided by Dr J. Carrington, Texas A&M University). Oligonucleotide primers for amplification of Nla were TEVNIA.N and TEVNIA.C (see Table 1 for the sequences of all primers used in this work). These two primers amplified the Nla open reading frame.

Fig. 1. Schematic diagram of the expression cassette pPRO1. The open box represents the Nla open reading frame. The shaded areas enlarged above show (as both nucleotide and amino acid sequence) the heptapeptide recognition sequence for the Nla proteolytic activity at both N- and C-termini of Nla; the engineered Sma I and Stu I cloning sites (underlined) for the in frame introduction of different genes; and start ATG and stop TGA codons. The Nla processing site between Gln and Gly is indicated as an open arrowhead. The sequence of the TEV 5’ untranslated region is also indicated with a black arrow upstream of Nla. Relevant unique restriction enzyme sites are indicated: Ba (Bam HI), Bg (Bgl II), Ec (Eco RI), Sa (Sal I), Sc (Sac I), Sm (Sma I), and St (Stu I).