Direct immunological identification of full-length cDNA clones for plant protein without gene fusion to *E. coli* protein

Kenzo Nakamura, Tsukaho Hattori and Tadashi Asahi

Laboratory of Biochemistry, Faculty of Agriculture, Nagoya University, Chikusa, Nagoya 464, Japan

Received 10 January 1986

By immunological screening of a cDNA library constructed from potato tuber poly(A)$^+$ RNA and *Escherichia coli* expression vector pUC8 by the vector-primer and linker procedure of Okayama and Berg [(1982) Mol. Cell Biol. 2, 161–170], nearly full-length cDNA clones for patatin, a major protein of potato tuber, were identified. The cDNA carrying part of the 5′-noncoding region of the patatin mRNA, in addition to entire coding and 3′-noncoding regions, expressed prepatatin in *E. coli* cells by translational initiation inside cDNA. These results suggest that nearly full-length cDNA clones with entire coding region can be identified directly by immunological screening without gene fusion to *E. coli* proteins at least for some plant mRNAs.

cDNA *Immunological screening* *Direct expression* *Patatin* *(Potato tuber)*

1. INTRODUCTION

Recent advances in the technology of molecular cloning of cDNAs have made it possible to construct cDNA libraries from various poly(A)$^+$ RNA sources, the major difficulty lying in identification of the cDNA clone of particular interest. Although sequence analysis of the cDNA insert and in vitro translation of hybridization-selected mRNA are straightforward methods for the final identification of cDNA clones, such assays are not usually suited to screening large cDNA libraries. One method developed for this purpose is the immunological screening of cDNA 'expression' libraries [2–4]. In this method, the cDNA preparation is cloned into 'expression vectors' that promote expression of cDNA in *E. coli*, and colonies expressing an antigenic structure are identified by an immunological assay procedure using specific antibody. By the improved technique for in situ colony immunological assay [3], one can easily screen many thousands of colonies in a relatively short time.

A crucial point of this procedure is the expression of antigenic structure in *E. coli*. Since the mechanism of translational initiation differs for *E. coli* and eukaryotic mRNAs, it is generally believed that gene fusion of a cDNA coding sequence to an *E. coli* structural gene sequence on an expression vector to produce a fusion protein in *E. coli* cells is necessary for this purpose [2–4]. Usually, the double-strand cDNA synthesized by the conventional S_1 nuclease method is used to construct cDNA expression libraries, since this method often removes the 5′-side of the mRNA sequence from the cDNA which in turn facilitates the fusion of bacterial and eukaryotic coding sequences. The major drawback of this approach is that the cDNA clones identified by this method lack the important N′-terminal coding and 5′-noncoding sequence of the mRNAs. To obtain full-length cDNAs, it is necessary to repeat screening of the library by colony hybridization assay.

The Shine-Dalgarno sequence, which is required for translational initiation in *E. coli*, is quite variable among various mRNAs in their length and

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

degree of complementarity to 16 S rRNA, and position relative to the ATG codon [5]. It seems likely that even eukaryotic mRNA sequences can be directly translated to a polypeptide with the number of copies sufficient to be detected by highly sensitive in situ colony radioimmunoassay [3], provided adequate amounts of mRNAs are produced in a bacterial cell. To test this hypothesis, we applied the vector-primer and linker procedure of Okayama and Berg [1] for the efficient synthesis of full-length cDNAs to restriction enzyme fragments of the expression vector pUC8 [6] for the construction of cDNA expression library from potato tuber poly(A)$^+$ RNA. We found that all of the 6 colonies identified to be immunologically positive by colony radioimmunoassay of the library with antiserum against patatin, one of the major tuber specific proteins of potato [7], contained nearly full-length patatin cDNAs.

2. MATERIALS AND METHODS

2.1. *Plant material and poly(A)$^+$ RNA*

Poly(A)$^+$ RNA was prepared from developing tubers of potato (*Solanum tuberosum* L. cv. Wheeler) as described [8], and size-fractionated on a 5–20% sucrose density gradient centrifugation for 16.5 h at 27000 rpm in a Hitachi RPS40T rotor.

2.2. *Preparation of vector-primer and tailed-linker*

The expression vector pUC8 carrying promotor and the amino-terminal region of the *E. coli lacZ* gene [6] were used for construction of the cDNA expression library. Homopolymer tails averaging 65 (dT) residues were added to the 3′-termini of *Pst*I-cleaved pUC8 in a reaction mixture (80 μl) containing 24 μg *Pst*I-cut pUC8, 140 mM sodium cacodylate (pH 7.0), 0.1 mM dithiothreitol (DTT), 40 μM dTTP, 2 mM $MnCl_2$ and 60 U of the terminal deoxynucleotidyl transferase (Takara Shuzo, Kyoto) at 30°C for 7 min. The (dT) tail on one strand was removed by digestion with *Hin*cII, and the large DNA fragment purified by polyacrylamide gel electrophoresis.

The oligo(dG)-tailed linker fragment was prepared by adding homopolymer tails averaging 25 (dG) residues to the 3′-termini of *Pst*I-cut pUC8 in 400 μl reaction mixture containing 250 μg *Pst*I-cut pUC8, 140 mM sodium cacodylate (pH 7.0), 0.1 mM DTT, 73 μM dGTP, 1 mM $CoCl_2$, and 100 U terminal transferase at 37°C for 4 min. The DNA was digested with *Eco*RI and the (dG)-tailed short (~30 bp) *Eco*RI-*Pst*I fragment purified by polyacrylamide gel electrophoresis.

2.3. *Construction of cDNA library*

The first strand of cDNA was synthesized from 2.1 μg size-fractionated poly(A)$^+$ RNA and 1.3 μg vector-primer DNA. After the homopolymer tails of about 15 – 20 (dC) residues were added to the 3′-termini of cDNA, the DNA was digested with *Eco*RI to remove the (dC) tail added to the *Hin*cII termini. One-fortieth of the resulting plasmid-cDNA, corresponding to 24 ng vector primer, was annealed with 2.5 ng of the (dG)-tailed linker fragment, and the second strand of cDNA synthesized using RNase H and *E. coli* DNA polymerase after cyclization with *E. coli* DNA ligase [1]. When introduced into *E. coli* JM83 [6] cells by the transformation method of Hanahan [9], these DNAs gave about 6000 ampicillin-resistant colonies.

2.4. *Other methods*

In situ colony radioimmunoassay was carried out as described by Helfman et al. [3] except that ^{125}I-protein A (Amersham, 30 mCi/mg; 1 Ci = 37 GBq) was used instead of second antibody. In vitro translation in a wheat-germ system, immunoprecipitation, SDS-polyacrylamide gel electrophoresis and fluorography were carried out as in [10]. *E. coli* cells were pulse labelled with [^{35}S]methionine, and fractionated as described [11]. Methods for nick translation of the cDNA insert, Northern blot hybridization, and hybridization selection of mRNAs were according to [12]. DNA sequencing was carried out by the M13-derived dideoxynucleotide chain-termination method [13].

3. RESULTS AND DISCUSSION

A cDNA expression library was constructed from potato tuber poly(A)$^+$ RNA and *E. coli* expression vector pUC8 by the vector-primer method of Okayama and Berg [1] as described in section 2. We screened about 550 colonies from the library by colony hybridization probed with ^{32}P-labelled

cDNA synthesized from potato tuber poly(A)$^+$ RNA by the oligo(dT) primer method [14]. Among those which gave strongly positive signals, 60 colonies were further screened by in situ colony radioimmunoassay [3] with anti-patatin serum and ^{125}I-protein A.

As shown in fig.1, 5 colonies consistently reacted with anti-patatin serum. The plasmid DNAs were isolated from the 5 colonies and another colony which gave a positive signal in one experiment but not in that shown in fig.1. The reason why one clone failed to react with anti-patatin serum in this experiment is not clear. All the 6 plasmid DNAs carried cDNA inserts of 1.4–1.52 kb which carried internal *Hin*dIII and *Bam*HI sites in common at almost identical locations from the *Pst*I site. Further restriction enzyme analysis, however, revealed that not all of these plasmids are identical. Restriction enzyme maps of two of the representative clones are shown in fig.2. Plasmids pKEN8-37 and pKEN8-42 carried cDNA inserts of 1.4 and 1.52 kb, respectively. Alignment of internal *Hin*dIII and *Bam*HI sites of these two and other cDNAs indicates that the difference in their length is primarily due to the difference in their 3′-end. The *Kpn*I site is absent in pKEN8-42.

The restriction enzyme maps of these plasmids are very similar to those of cDNA clones of patatin

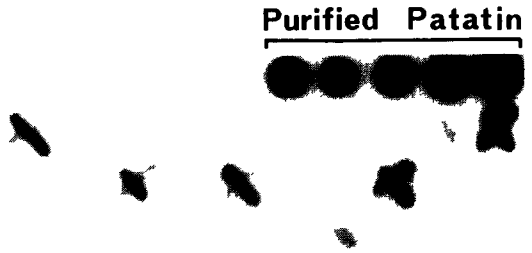

Fig.1. In situ colony radioimmunoassay of cDNA clones with anti-patatin serum. 60 colonies from the library were streaked on a nitrocellulose filter placed on an L broth plate containing ampicillin (50 μg/ml). After cells were grown overnight at 37°C, in situ colony radioimmunoassay with anti-patatin serum (1:50 dilution) and ^{125}I-protein A was carried out. Six colonies which gave positive signals in another assay were marked by crossing the streak. In this experiment, 5 showed positive reactions.

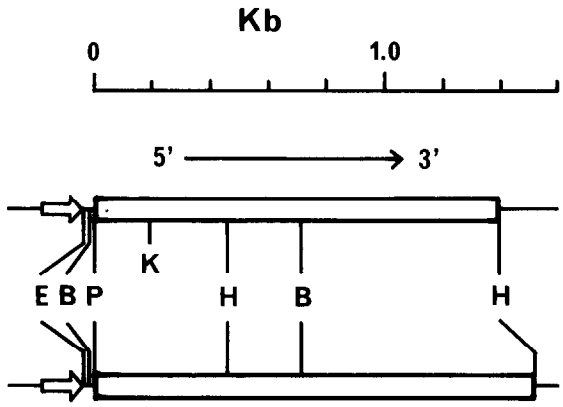

Fig.2. Restriction enzyme maps of two representative patatin cDNA clones identified by immunological screening. The thin lines represent sequences derived from pUC8, and arrows denote the approximate position and direction of the *lac* control region. Homopolymer tails of (dG:dC) and (dA:dT) are included as part of cDNA inserts (▭). E, *Eco*RI; B, *Bam*HI; P, *Pst*I; K, *Kpn*I; H, *Hin*dIII.

from a different potato cultivar reported by Mignery et al. [15]. They also reported two closely related patatin cDNA clones. Indeed, both pKEN8-37 and pKEN8-42 DNAs specifically hybridization-selected mRNAs coding for polypeptides with the same electrophoretic mobility as that of prepatatin, an in vitro precursor of patatin, and these polypeptides were immunoprecipitated specifically with anti-patatin serum (not shown). Northern blot analysis of tuber poly(A)$^+$ RNA with a ^{32}P-labelled cDNA insert of pKEN8-42 indicated the size of patatin mRNA to be approx. 1450 nucleotides, which is very close to that reported by Mignery et al. [15]. These results indicate that all 6 cDNA clones are essentially full-length cDNA clones of heterogeneous patatin mRNAs.

We determined the nucleotide sequence of the 5′-proximal one-third of the pKEN8-37 cDNA insert. Since about 30 (dG) residues following the *Pst*I site hindered the sequencing of the region further downstream, we used Bal31 nuclease to delete the (dG) tails which also removed 3–4 nucleotides from the cDNA portion. The sequence showed about 96% homology with those of patatin cDNA clones pGM01 and pGM203 reported by Mignery et al. [15]. The difference could be due to differences in potato cultivars used and/or

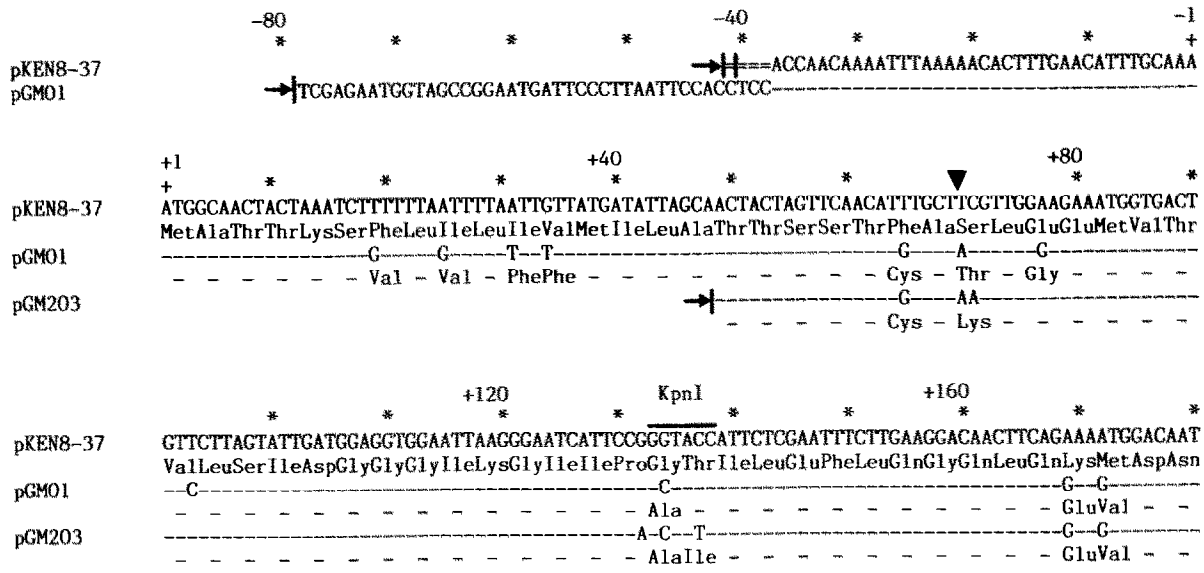

Fig.3. The 5′-proximal nucleotide sequence of pKEN8-37 cDNA compared with those of pGM01 and pGM203 [15]. The cDNA insert of each clone begins at (→|). Nucleotides removed from the 5′-end of pKEN8-37 cDNA (see text) are indicated by (=). Sequences are compared by indicating homology with dashed lines, and differences in nucleotide or amino acid sequence of pGM01 and pGM203 from those of pKEN8-37 are indicated. The *Kpn*I cleavage site sequence in pKEN8-37 is absent in other cDNAs. The amino terminus of mature patatin is indicated by (▼).

microheterogeneity of patatin mRNAs. Fig.3 shows the 5′-proximal nucleotide sequence of pKEN8-37 compared with those of pGM01 and pGM203. The cDNA insert of pKEN8-37 contained an at least 37 bp 5′-noncoding region, which is about 40 bp shorter than that of pGM01.

To determine whether the patatin-related antigen expressed in *E. coli* cells harboring pKEN8-37 or pKEN8-42 is the fusion protein with the N′-terminal portion of the *lac*Z coded by the vector, we examined the pulse-labelled cellular proteins by immunoprecipitation with anti-patatin serum. If the patatin-related antigen is translated as a fusion protein, the size of the fusion protein should be expected to be about 35 amino acids larger than prepatatin, since the 5′-noncoding region of pKEN8-37 does not contain termination codons in the same reading frame as that of prepatatin. As shown in fig.4, anti-patatin serum predominantly precipitated a polypeptide with the same electrophoretic mobility as that of prepatatin (lane 5), from both soluble (lane 3) and membrane (lane 1) fractions of *E. coli* JM103 [6] harboring pKEN8-37. In the case of pKEN8-42, a polypeptide about 2 kDa smaller than prepatatin was also detected as a major band in the soluble fraction, in addition to a polypeptide with the same size as prepatatin (lane 2). These results indicate that these patatin-related antigens are not synthesized as fusion proteins, and suggest that translations are mainly initiated from ATG codons inside of these cDNAs to produce prepatatin itself.

Judging from the size and location of these patatin-related polypeptides, it is possible that the signal peptide is not cleaved off from the prepatatin coded by pKEN8-37, while some of the signal peptide is cleaved off from the prepatatin coded by pKEN8-42. We have not determined the nucleotide sequence of pKEN8-42. Since the N′-terminal amino acid of patatin species shows a high degree of heterogeneity [7] (fig.4), the signal peptide cleavage site sequence of these two prepatatin may be different and cleaved by *E. coli* signal peptidase with different efficiency.

Although we first screened the library with a cDNA hybridization probe, the present results indicate that nearly full-length cDNA clones with entire coding regions can be identified directly by immunological screening of the cDNA expression library constructed by the vector-primer method

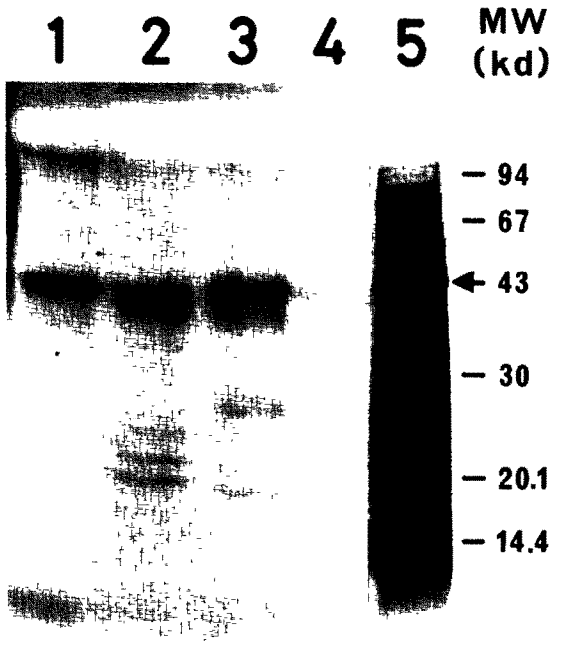

Fig.4. Synthesis of patatin-related antigen in *E. coli*. *E. coli* JM103 harboring pUC8 (lane 4), pKEN8-37 (lanes 1,3), or pKEN8-42 (lane 2) were pulse-labelled with [^{35}S]methionine after induction for *lac* expression by IPTG, and the soluble (lanes 2–4) or membrane (lane 1) fractions were subjected to immunoprecipitation with anti-patatin serum after solubilization in 2% SDS. Lane 5 shows total translation products of poly(A)$^+$ RNA from potato tuber translated in a wheat-germ extract in the presence of [^{35}S]methionine. The position of prepatatin is indicated by an arrowhead (◄).

without fusing the cDNA coding sequence to the *E. coli* protein coding sequence. Since all 6 immuno-positive colonies contained nearly full-length patatin cDNAs, it seems that this method rather preferentially identifies those cDNA clones which have longer inserts. It is not clear how this method can be applied in general. However, we have successfully used the same approach for direct immunological identification of nearly full-length cDNA clones for two other tuber specific proteins of potato (Nakamura, K., unpublished), a major storage protein (sporamin) [8] and catalase (Sakajo, S. et al., in preparation) from sweet potato tuberous roots. This method considerably simplifies the establishment of a cDNA expression library and the cloning of nearly full-length cDNAs.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Agricultural Chemical Research Foundation, Mishima Kaiun Science Foundation and Plantech Research Institute.

REFERENCES

[1] Okayama, H. and Berg, P. (1982) Mol. Cell Biol. 2, 161–170.
[2] Broome, S. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA 75, 2746–2749.
[3] Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P. and Hughes, S.H. (1983) Proc. Natl. Acad. Sci. USA 80, 31–35.
[4] Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
[5] Steitz, J.A. (1979) in: Biological Regulation and Development (Goldberger, R.F. ed.) vol.1, pp.349–399, Plenum, New York.
[6] Vieira, J. and Messing, J. (1982) Gene 19, 259–268.
[7] Park, W.D., Blackwood, C., Mignery, G.A., Hermodson, M.A. and Lister, R.M. (1983) Plant Physiol. 71, 156–160.
[8] Hattori, T., Nakagawa, T., Maeshima, M., Nakamura, K. and Asahi, T. (1985) Plant Mol. Biol., in press.
[9] Hanahan, D. (1983) J. Mol. Biol. 166, 557–580.
[10] Hattori, T., Sakajo, S., Iwasaki, Y. and Asahi, T. (1982) Biochem. Biophys. Res. Commun. 113, 235–240.
[11] Nakamura, K., Masui, Y. and Inouye, M. (1982) J. Mol. Appl. Genet. 1, 289–299.
[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] Messing, J. (1983) Methods Enzymol. 101, 20–78.
[14] St. John, T.P. and Davis, R.W. (1979) Cell 16, 443–452.
[15] Mignery, G.A., Pikaard, C.S., Hannapel, D.J. and Park, W.D. (1984) Nucleic Acids Res. 12, 7987–8000.