Acta Biochimica Polonica

Vol. 41 No. 2/1994

QUARTERLY

Expression of *Lupinus luteus* nodulin-45 gene in *E. coli* cells. Mutagenesis of coding sequence by PCR method*

Michał M. Sikorski and Marek Rożek

Institute of Bioorganic Chemistry, Polish Academy of Sciences, Z. Noskowskiego 12/14, 61-704 Poznań, Poland

Key words: expression, mutagenesis, nodulin-45, PCR, yellow lupin

The nodule specific cDNA clone — pLN-13 [1] from yellow lupin encodes a single polypeptide of molecular mass 45 kDa (called nodulin-45; N-45[1]). Figure 1 shows a full length cDNA copy of pLN-13. A deduced amino-acid sequence reveals that this protein is highly hydrophilic with putative N-terminal signal peptide and two asparagine-linked glycosylation sites. The N-45 cDNA clone (1399 bp) represents one of the abundant sequences of lupin cDNA library. A coding region of 1200 bp is located between nucleotides #41 and #1240 of the full copy of pLN-13 cDNA clone (subcloned into pBluescript SK at *Pst*I of multiple cloning site). The function and subcellular localization of nodulin-45 still remains unclear.

To prepare N-45 protein for functional studies and determination of subcellular localization we decided to express its cDNA clone in *E. coli* cells in the pET-3a expression vector [2]. In this system, synthesis of large amounts of foreign gene products is directed by phage T7 gene 10 promoter, which uses T7 RNA polymerase. It is well known that the polymerase transcribes the gene 10 promoter very efficiently.

The nodulin-45 cDNA coding fragment of 1200 bp has been amplified by PCR. As a target for the amplification reaction, a full cDNA copy (1399 bp) inserted to pBluescript SK at *Pst*I site was used. The oligodeoxynucleotide primers (N & C) were synthesized to amplify coding fragment of N-45 cDNA clone in reaction with Taq DNA polymerase. The N-primer containing the sequence recognized by *Nde*I matches 3'-end of n45 cDNA clone. The C-primer contains the sequence recognized by *Bam*HI and two stop codons. It matches the complementary strand of 5'-end of this clone.

Since *Bam*HI was used for insertion of the amplified sequence into the pET-3a expression vector it was necessary to introduce mutation at the internal *Bam*HI site of cDNA clone. Two additional (internal primers: a and b, Fig. 2) which matched the internal *Bam*HI sequence were mutated at the position corresponding to nucleotide #385 of the n45 clone ($G_{385} \rightarrow T_{385}$). This mutation (in the triplet encoding arginine: AGG → AGT) did not change the amino acid encoded by new triplet (AGT).

The n45 cDNA coding sequence with mutated nucleotide at the position #385 was prepared by two independent ways (variant 1& 2):

– 1. Amplification of n45 coding sequence (1200 bp) from pBluescript SK (CC form) carrying a full copy of cDNA clone (pBluescript/n45) using two pairs of primers: external (N & C) and internal (a & b). The molar ratio of target DNA to external and internal primers was 1:10000:10, respectively. PCR products and their restriction fragments after digestion

*This work was partially supported by the Polish Academy of Sciences within the project No. 6 6281 92 03.

[1] Abbreviations: bp, base pair; N-45, yellow lupin nodulin-45; n45, cDNA coding fragment for nodulin 45; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium salt of dodecyl sulphate.

```
   1    A AAC CAT GGC TTC TTT TTC ATT TGC TTT GCA AAG AGT GTT ATG GAT AAG TCA AAG      55
   1                                                      Met Asp Lys Ser Lys       5

  56  ATT TGT GTT TCC TTT ATT CTT CTT TTC ACT TTC CTC TTA TCT TCT CAT GTC TTG GCT     112
   6  Ile Cys Val Ser Phe Ile Leu Leu Phe Thr Phe Leu Leu Ser Ser His Val Leu Ala    24

 113  TAT GAA TTG CCT GAG ACT TTA CTC GGT GAA TCT TCC TTT GAT TCT GTC ACC CAA AAA     169
  25  Tyr Glu Leu Pro Glu Thr Leu Leu Gly Glu Ser Ser Phe Asp Ser Val Thr Gln Lys    43

 170  GCT ATA GAT ATC GGA TTT CCA TCT CCA CCG GAT GTG GCG ACA GAT TTC GAG AGA CAG     226
  44  Ala Ile Asp Ile Gly Phe Pro Ser Pro Pro Asp Val Ala Thr Asp Phe Glu Arg Gln    62

 227  GAA GTT GTT GAG AAA CAA GAG GAT CAA CTA CAC TCT CAC AAT GAC AAA GGA AAT GTG     283
  63  Glu Val Val Glu Lys Gln Glu Asp Gln Leu His Ser His Asn Asp Lys Gly Asn Val    81

 284  GTC AAG AAT CCA AAA CCA ACA TCA TCA AAG AAT TTG GAA ATG GAC CAC AAT GGC GTT     340
  82  Val Lys Asn Pro Lys Pro Thr Ser Ser Lys Asn Leu Glu Met Asp His Asn Gly Val    100

 341  TCT GGC ATG TTC AGT TTA AGG CAC TCT GCG GGA CAT GGT TCC AGG ATC CCA ACA TTG     397
 101  Ser Gly Met Phe Ser Leu Arg His Ser Ala Gly His Gly Ser Arg Ile Pro Thr Leu    119

 398  GAA GAA AAA GAG AAT CAA CGC TAT GAG GAA GGA AAG GTG GAA CGA AAT CCA ACA TCA     454
 120  Glu Glu Lys Glu Asn Gln Arg Tyr Glu Glu Gly Lys Val Glu Arg Asn Pro Thr Ser    138

 455  TCG TAT GAT TTG GAA AGA AAG AGG AGA AGT ACT GAT TCT AGC ATG CAC AGA TTT AGC     511
 139  Ser Tyr Asp Leu Glu Arg Lys Arg Arg Ser Thr Asp Ser Ser Met His Arg Phe Ser    157

 512  CAG ACA CAT GGT TAC AAG ACC ACA TCA TCT GAA GAA ACA GGC AAG AGT GAT GTT TCA     568
 158  Gln Thr His Gly Tyr Lys Thr Thr Ser Ser Glu Glu Thr Gly Lys Ser Asp Val Ser    176

 569  AAA GAT ATT GGA GAG GCT CAT CAG TTG GGA CTG GAC TGG AAT ATA GGA TTT GGA AGT     625
 177  Lys Asp Ile Gly Glu Ala His Gln Leu Gly Leu Asp Trp Asn Ile Gly Phe Gly Ser    195

 626  GGT GGC GGT GGT AGG TCA AAT CGC GGT GGC GGC TTT CCT GGT ATC GGT GGC AAC TTT     682
 196  Gly Gly Gly Gly Arg Ser Asn Arg Gly Gly Gly Phe Pro Gly Ile Gly Gly Asn Phe    214

 683  CCT GGT ATC GGT GGT GAC ATC GGC AGT GGT TGG GGC GGT GAC TTC ATG AGT ACC GAG     739
 215  Pro Gly Ile Gly Gly Asp Ile Gly Ser Gly Trp Gly Gly Asp Phe Met Ser Thr Glu    233

 740  AAG CAA AAC AAT GAC GAG GGA AAG GTC TTT AAG GAA CGA AAT CCA ACA TCA TCG TAT     796
 234  Lys Gln Asn Asn Asp Glu Gly Lys Val Phe Lys Glu Arg Asn Pro Thr Ser Ser Tyr    252

 797  GAT TTG GAA AGA AAG ATG AGA AGT GGT TCT AGC ATG CCC AGT TTT AGC CAC TTT GTG     853
 253  Asp Leu Glu Arg Lys Met Arg Ser Gly Ser Ser Met Pro Ser Phe Ser His Phe Val    271

 854  AGA CAA GGT TAC AAA ACT CCA TCA TTG GAA GAA ACA GGC AAG AGT GAT GTT TCA AAA     910
 272  Arg Gln Gly Tyr Lys Thr Pro Ser Leu Glu Glu Thr Gly Lys Ser Asp Val Ser Lys    290

 911  GAT ACT GGA GAG GCT CAT CAG TTG GGA GTG GAC TGG AAA GTA GAA ATT GGA AGT GGT     967
 291  Asp Thr Gly Glu Ala His Gln Leu Gly Val Asp Trp Lys Val Glu Ile Gly Ser Gly    309

 968  GGC GGT CGT AGG TCA AAT CGC GGT GGC GGC TTT CCT GGT ATC GAT GGC GGC TTT CCT    1024
 310  Gly Gly Arg Arg Ser Asn Arg Gly Gly Gly Phe Pro Gly Ile Asp Gly Gly Phe Pro    328

1025  GGT ATC GGT GGT GAC ATC GGC AAT GGT GGT TGG GTC GGT GAC TTC ACG AGT ACC GAA    1081
 329  Gly Ile Gly Gly Asp Ile Gly Asn Gly Gly Trp Val Gly Asp Phe Thr Ser Thr Glu    347

1082  CCA ATG AAT TAC AAA GAT GCA AGT ATA ATG AGG CAT CAG AAT ATG AAG AAA TTT TCA    1138
 348  Pro Met Asn Tyr Lys Asp Ala Ser Ile Met Arg His Gln Asn Met Lys Lys Phe Ser    366

1139  GAA TTG GAT CGT GAG TCT CAA GTA GAC AGA AAA TTT GGG ATG AAG AGT GGT GGT AAA    1195
 367  Glu Leu Asp Arg Glu Ser Gln Val Asp Arg Lys Phe Gly Met Lys Ser Gly Gly Lys    385

1196  ATC TAC TAT GGT GGT GTC CCA AGA ATG GAA GAG GAT AAG AAG AAT TAG TGG TCG TTT    1252
 386  Ile Tyr Tyr Gly Gly Val Pro Arg Met Glu Glu Asp Lys Lys Asn ***              401

1253  AGG TTT TTA TGA ACC ATA TAG AAC CAT TAT GTC TTC GCC AAA TAA TAC CCT AAA TAA    1309

1310  GAT GCA TGG TGC ATC TTA CAA TAT GAT GAC ACT AAA TAT GTA CTT ATT CTA TGT ATT    1366

1367  TCT TGT TCC ATT ACT ATA GTA AGT TTT ATA AAA                                   1399
```

Fig. 1. The nucleotide sequence of full copy cDNA clone of pLN-13.
This is a corrected sequence of pLN-13 published earlier [1] which was send to GeneBank and registered under the accession No.: X77044.

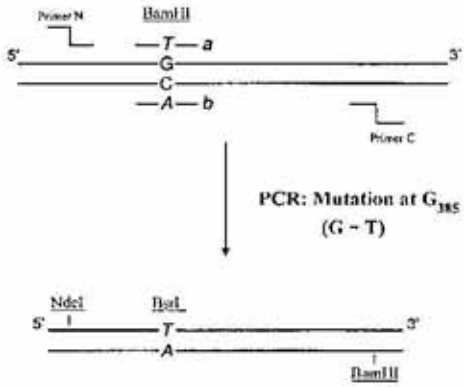

Fig. 2. Scheme of point mutation in N-45 coding sequence at the position of G385:
G → T, *Bam*HI site →*Bsr*I site; a, b, internal PCR primers causing mutation at position #385.

with *Bam*HI were analyzed by 1% agarose gel electrophoresis [3] (Fig. 3, lines 2, 3). Two PCR products of 1200 bp — with and without internal mutation were observed. The product without internal mutation digested with *Bam*HI yielded two fragments: 400 and 800 bp.

–2. Amplification of n45 coding sequence (1200 bp) from the linearized form of pBluescript/n45 (with *Bam*HI). This approach was carried out to increase the yield of the product with internal mutation. PCR products and their restriction fragments after digestion with *Bam*HI were analyzed on agarose gel (Fig. 3, lines 5, 6).

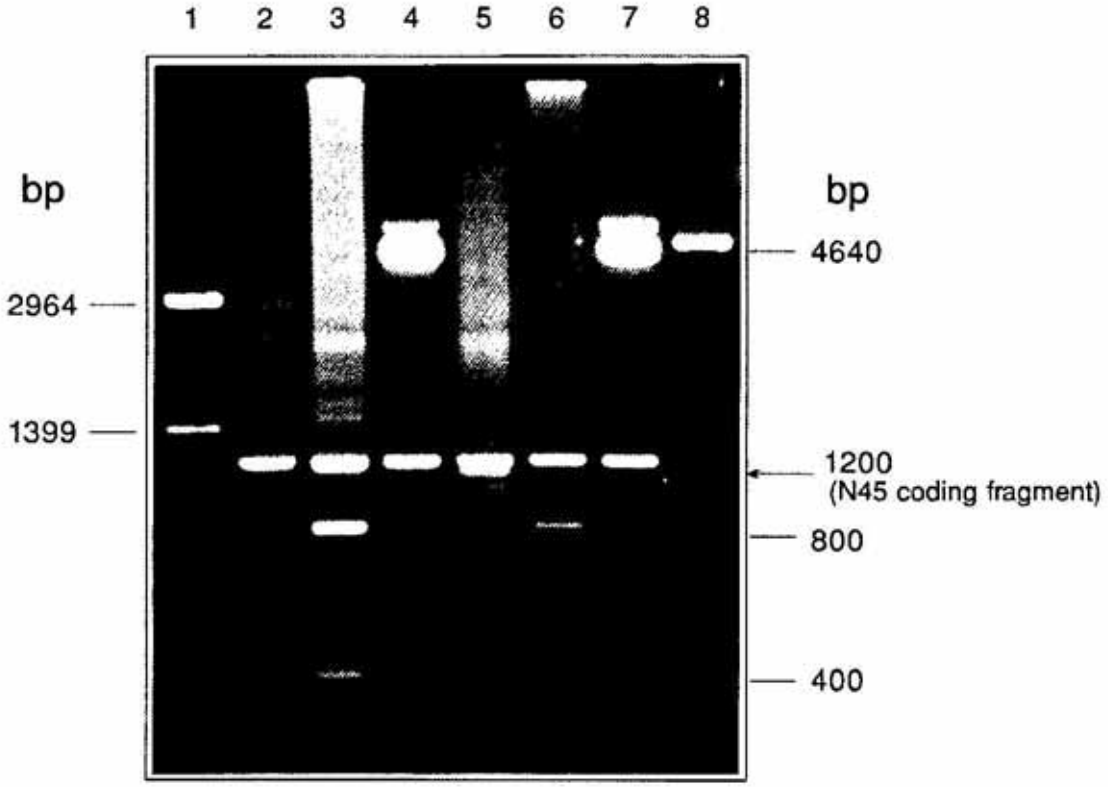

Fig. 3. Restriction analysis of cDNA coding fragment of N-45.
1, pBluescript/n45 digested with *Pst*I: 2964 bp — linearized pBluescript SK, 1399 bp — full copy of cDNA clone of N-45; 2, products of PCR amplification (1200 bp) with two pairs of primers, var. 1; 3, *Bam*HI restriction fragments of PCR products, var. 1: 1200 bp — cDNA coding fragment of N-45 with mutation at 385 nt, 800, 400 bp — *Bam*HI restriction fragments of amplified cDNA without mutation; 4, pET-3a/n45 digested with *Bam*HI and *Nde*I: 4640 bp — pET-3a, 1200 bp — n45; 5, products of PCR amplification (1200 bp) with two pairs of primers, var. 2; 6, *Bam*HI restriction fragments of PCR products, var. 2: 1200 bp — cDNA coding fragment of N-45 with mutation at 385 nt, 800, 400bp — *Bam*HI restriction fragments of amplified cDNA without mutation; 7, pET-3a/n45 digested with *Bam*HI and *Nde*I: 4640 bp — pET-3a, 1200 bp — n45; 8, pET-3a itself, linearized (with *Nde*I or *Bam*HI); (ad. 4 & 8: pET-3a/n45 was amplified in *E. coli* cells, strain HMS174 for isolation of pure plasmid for sequence analysis).

Products of PCR amplifications of the n45-coding fragment with mutation at the #385 nt and pET-3a were digested with *Nde*I and *Bam*HI and then ligated (pET3a/n45, Fig. 4).

E. coli cells, strain HMS174 [2] were transformed with pET-3a/n45 construct. Growing plasmid was purified on Qiagen column for its nucleotide sequence analysis to determine the right clone (sequence correspoding to the cDNA coding fragment mutated at the position #385 nt). The expression of pET-3a carrying correct n45 coding sequence was performed in *E. coli* cells, strain BL21(DE3)pLysS with T7 lysozyme activity [2]. T7 RNA polymerase gene was induced by addition to the culture of 0.5 mM IPTG. The products of gene expression were analyzed by 15% PAGE/SDS (Fig. 5) according to Laemmli [4]. The identity between expressed polypeptide and amino-acid sequence of nodulin 45 protein deduced from nucleotide sequence of cDNA clone (Fig. 1) was supported by microsequencing analysis of 12-

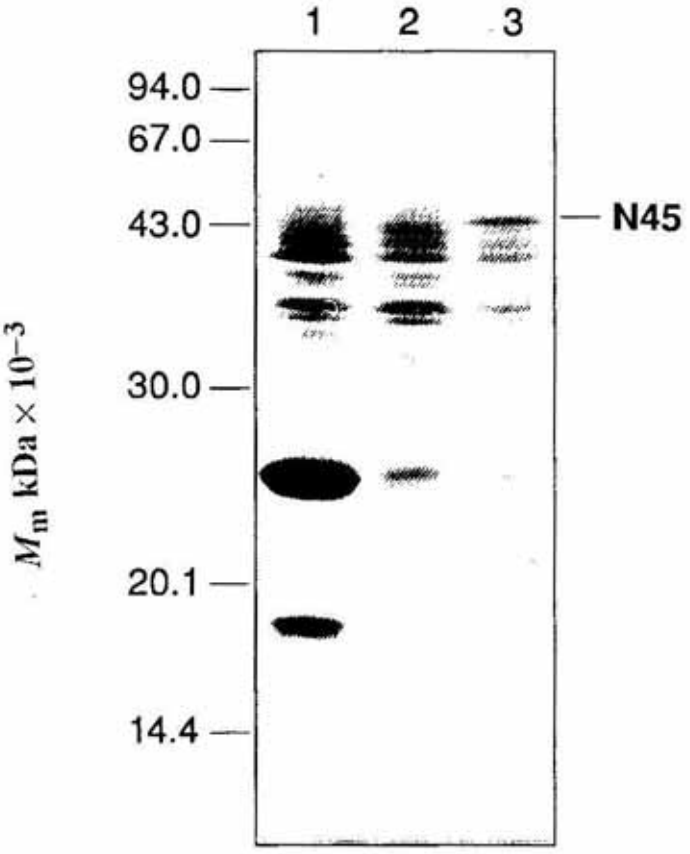

Fig. 5. 15% PAGE/SDS of total proteins of bacterial lysates:
1, protein extract of non-transformed cells; 2, protein extract of transformed, not induced cells; 3, protein extract of IPTG induced cells. Expressed N-45 is indicated (line 3).

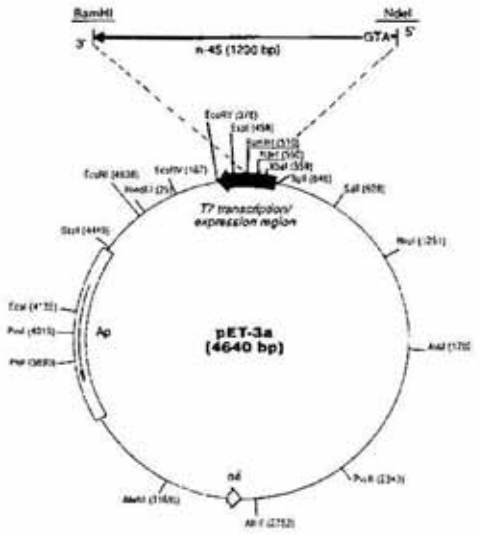

Fig. 4. Restriction map of pET-3a expression vector.
Position and orientation of lupin N-45 coding fragment (1200 bp) of cDNA clone are indicated.

N-terminal amino-acid residues of recombinant protein (electrobloted into PVDF membrane).

The applied PCR amplification method resulted in obtaining the designed point mutations within the target sequence when the pair of internal primers was used together with external primers. This method can be used to generate any point mutation at the level of nucleotides and aminoacids. Genetically engineered proteins obtained in expression system in bacterial cells can be further studied either by X-ray crystallography or functional assays.

We thank Professor A.B. Legocki in whose laboratory this work was performed for stimulating and helpful discussion and critical reading of the manuscript. We are very grateful to Ms Alina Akuszewska for technical assistance.

REFERENCES

1. Szczyglowski, K., Boroń, L., Szybiak-Stróżycka, U. & Legocki, A.B. (1989) *Plant Science* **65**, 87–95.
2. Studier, F.W., Rosenberg, A.H., Dunn, J.J. & Dubendorff, J.W. (1990) *Methods in Enzymology* **185**, 60–89.
3. Sambrook, J., Fritsh, E.F. & Maniatis, T. (1989) *Molecular cloning, a laboratory manual*. Cold Spring Harbor Lab.
4. Laemmli, U.K. (1970) *Nature (London)* **227**, 680–685.