Volume 262, number 1, 29–32 FEBS 08206 March 1990

A gene from pea (*Pisum sativum* L.) with homology to metallothionein genes

I. Marta Evans, Laurence N. Gatehouse, John A. Gatehouse, Nigel J. Robinson and Ronald R.D. Croy

Plant Molecular Biology Research Group, Department of Biological Sciences, University of Durham, South Road, Durham DH1 3LE, England

Received 17 November 1989; revised version received 13 December 1989

While searching for 'organ-specific' genes in pea (*Pisum sativum* L.) we have isolated a gene (designated $PsMT_A$) which has an ORF encoding a predicted protein with some similarity to metallothioneins (MTs). The $PsMT_A$ transcript is abundant in roots which have not been exposed to elevated concentrations of trace metals.

$PsMT_A$; 5′ Flanking sequence; Metal metabolism; (*Pisum sativum*)

1. INTRODUCTION

Early reports of MT-like proteins in plants exposed to excess metals have been succeeded by reports of poly(γ-glutamylcysteinyl)glycine ($(\gamma EC)_nG$) [1–4], a metal-induced metal-binding polypeptide designated class III MT, which is not encoded by mRNA [5,6]. Many of the MT-like proteins isolated from higher plants are now thought to have been impure isolates of $(\gamma EC)_nG$ [3]. Probing Southern blots of plant DNA with fragments of class I MT genes has not identified any homologues. In this report we describe the isolation of a gene from garden pea which has an ORF encoding a protein with some similarity to metallothioneins. Confirmation of the identity of this gene requires the characterization of its putative product.

2. MATERIALS AND METHODS

2.1. *cDNA library production and screening*

Poly(A)$^+$ RNA was isolated from the roots of pea plants (*Pisum sativum* L., cv. Feltham First) [7] and used for the synthesis of cDNA [7]. cDNA was ligated to *Eco*RI linkers, inserted into *Eco*RI cut pUC18 and transformed into *E. coli* DH5α (BRL). Colonies (1152) were differentially screened and 64 clones selected containing sequences which were relatively more abundant in roots than in other organs (green leaf, etiolated leaf and cotyledon).

2.2. *Northern and Southern blotting*

Poly(A)$^+$ RNAs (2 μg/lane) prepared from developing cotyledon, root, green leaf and etiolated leaf were glyoxalated, analysed on a 1.5% agarose gel [8], and Northern blot analysis carried out as described previously [7]. The filter was washed to 0.1 × SSC, 0.1% SDS at 65°C. Selected cDNA probes were labelled by random oligo-priming [9]. Genomic DNA (10 μg/lane) prepared from pea leaves [10] was restricted and analyzed on a 0.7% agarose gel. Southern blots were washed to a final stringency of 0.1 × SSC, 0.1% SDS at 65°C.

2.3. *Genomic DNA library screening and sequencing*

A pea genomic DNA library in lambda EMBL3 was a gift of Anil Shirsat, University of Durham. Growth of the lambda phage and purification of the phage DNA were performed according to standard techniques [11]. The library was screened using selected cDNA probes. Hybridizing genomic fragments were restriction mapped, then subcloned in pUC18. Dideoxynucleotide DNA sequencing was carried out using [α-^{35}S]thio-dATP [12] or by modified dideoxynucleotide sequencing with Taq DNA polymerase on an automated DNA sequencer (Applied Biosystems model 370A).

3. RESULTS AND DISCUSSION

Five cDNA clones, out of the initial selection of 64, contained sequences which were most abundant in pea roots. The inserts from two of these clones, pPR179 and pPR705, cross-hybridised and were subsequently shown to be identical in sequence. Northern blots probed with pPR179, showed hybridization to a single, abundant transcript (~640 bases) in poly(A)$^+$ RNA from pea root (fig.1). Much weaker hybridisation was also observed to similar-sized transcripts in green and etiolated leaves, and to a smaller transcript in developing cotyledons. This probe also hybridised to multiple fragments in restricted pea genomic DNA (fig.2), suggesting a small multigene family.

A pea genomic DNA library was screened using the pPR179 probe. A lambda clone containing sequences hybridising to this probe was isolated, the DNA purified, restriction-mapped and the appropriate frag-

Correspondence address: I.M. Evans, Plant Molecular Biology Research Group, Department of Biological Sciences, University of Durham, South Road, Durham DH1 3LE, England

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

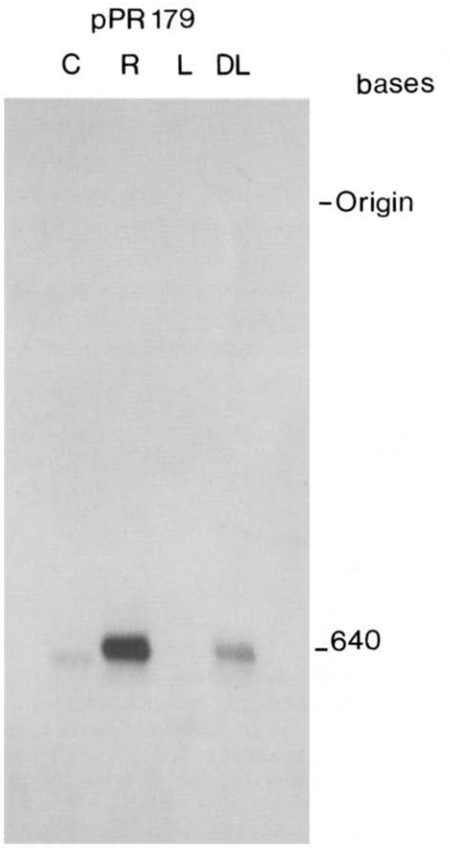


Fig.1. Northern blot analysis of poly(A)$^+$ RNA from four pea organs: developing cotyledon (C), root (R), green leaf (L) and etiolated leaf (DL), using the pea root-enhanced cDNA excised from pPR179 as a probe.

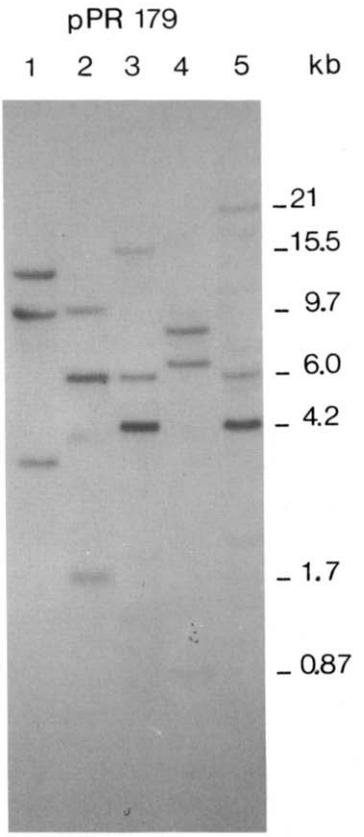


Fig.2. Southern blot of pea genomic DNA restricted with *Eco*RV (lane 1), *Eco*RI (lane 2), *Bam*HI (lane 3), *Hin*dIII (lane 4), *Bgl*II (lane 5) and hybridized with pea root-enhanced cDNA from pPR179 (sizes in kb).

ment subcloned. The sequence of the hybridising fragment is shown in fig.3. The coding strand of the gene was identified by using strand-specific probes from pPR179 in M13, mp18 and mp19 vectors [13] hybridised to Northern blots of pea root poly(A)$^+$ RNA, followed by sequencing of the probe showing positive hybridisation.

The cDNA and gene sequences differed only in the presence of an intron of 634 bp in the gene; when this was taken into account, the combined sequences encoded a putative polypeptide of 75 amino acids. The cDNA sequence extends from codon 12 of the ORF to 271 bp beyond the stop codon, but lacks a poly(A)$^+$ tail. The pea gene sequence contains ~800 bp of 5′-flanking sequence and 490 bp of 3′-flanking sequence as well as the complete coding sequence and the intron.

Computerized comparisons (Fast P software) of the predicted pea protein with the NBRF protein database selected exclusively MT sequences (monkey MT-1; fruit fly MT; horse MT-1B; horse MT-1A; Chinese hamster MT II; green monkey MT II; human and monkey MT-2; rat MT-II; mouse MT-II; *Neurospora crassa* MT) as the 10 best matches. By contrast the wheat E_c protein, a higher plant class II MT [14,15], did not show significant homology to sequences in the database. Alignment (2 breaks in the pea sequence) of the predicted amino acid sequence encoded by the pea gene with the 26 amino acids of class I MT from *N. crassa* reveals 14 identical amino acids. The probability that the detected homology has arisen at random on the basis of the defined amino acid compositions is less than 0.00025. The *z*-value for the comparison is 4.7 (RDF comparison, FAST P software). The pea gene was designated $PsMT_A$.

Matrix comparisons of the amino acid sequence of the predicted pea protein and class I MT from *N. crassa* identify two regions of homology at amino acids 4–18, and 61–74, in the predicted pea protein, designated as domains 1 and 2, respectively (fig.4). A feature of both domains is the presence of Cys-X-Cys motifs, where X is an amino acid other than Cys, a constraint of metal-binding and a characteristic of MTs [6]. These two domains are separated by a central region which contains no Cys and is dissimilar to class I MTs. Comparison of $PsMT_A$ with MT genes shows the location of the intron to be most similar to the *Neurospora crassa* MT gene [16] (after the second and first bases of codon 17 and codon 18, respectively).

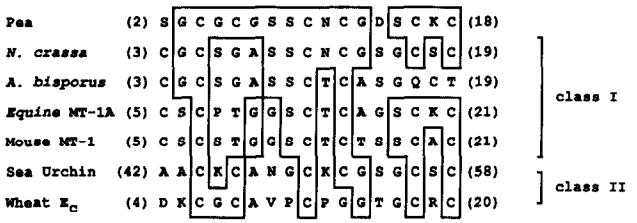
```
.....(-806)5'AACCA
AGTTTCTTGTTCTTCAAACATTTTCTAATTCTTCAGCTTAGAGTTTCCTTTCATAATTAATGTAAGAATTGACCCTAGTAAGAAAGTCACTCATACCCTT
TCCTTCTTCTAGCCCTAAATTTCCGTGAAACATATAGTCTTCCCTTAACCCTTTTTGGAACATCCAACATTGCAGCTTATCATATATTGATCCTACTTTG
GTGAAGCGTTATATGTACTCCCGGAGGATTTCCATTTTCATCTATTGGATTCCTATGAAGACAACTATAGTCTTGGATTTCAATTTTTTAATTAATCAAA
AATAAATTATCCCGTGACATTATTAAATACCATTTTGAAATAAAAAGTTTCTAAAAAAAAGTCCTTCGTCATTATCCATCATTTTTATTAAAGTATATTT
TATTATTTATTGAAACAAGTAATCATCCATTAGCGGAGAAATAGAGGAGTGAAGCATTTAACTTTTCCAAACGAAAGCGACGTAATCAACCTACATTTGA
CTTAGATTGGATTAAGCATGCAACAAATTAAAATTTAATCGCCATTGCAATTTGCACACCACAATAAGACGTGTGATGAAAACGATGATATCTACGTGGA
AATAATCCAAGGGTGGCCTTGTGGACCCATGCAACACAGGATGACAACACGTGGACGGTCAAGATTTCACCAATTATTCTCTCCCACCTTATAAATTGGG
GCACGCAACATCATTAAAAGACATCAATTGTAGTGAAGATAACAGCAACCAAGCAATTAATATCAATTGTTGTTTGCAAAAAATCTTAGGTTCTGAAAAT

ATG TCT GGA TGT GGT TGT GGA AGC AGT TGC AAC TGT GGT GAT AGC TGC AAGTAAGGATCCACCACCTTAATTCTTTGTTGTTTT
Met Ser Gly Cys Gly Cys Gly Ser Ser Cys Asn Cys Gly Asp Ser Cys Ly
            Domain 1                [
TCTGTATAATTTTTTCATTACAATTATTTGTATGTCTATTTTTAATCATATAGATGATTCTTTGGAGATTTTTTTAAATAATTTGTTTAGTTTTATCGCA

TCGAATAATATATGATCTGAGCATGAGAAAAATAAATTTAATATAGACGGATTGTTTTTTATAAATGAATTAGGCTGAATCTAAATTCTAAGACTATGAA

TATGGTTCATAATTCTATGTTAAATCATTTTGTGTAGTGAAATTGGGCAATTTTATGTGTAAACGCATAATTTTGAGGTTTAAAATAAGGATCGTGCTGT

CGCGATAGTTTAAGTGTCGATTGTAGTCGCGTAAAGGCTTTTCTGATTTCGGTTGGTTTAAGTGTTGATTGCAATCGTGTAAAGAGTTTTCTAATTTCGG

TTGGTTTAGGTGCGATTGCAGTCGGGTAAAGATTTTCGTGATTGTCGTCGTTGCGGTGTGAATTAATCACAATTTCTTCTTTATCATAAAAACGTTGAAT

AACATATCGATATCGATTTGAAAACCTTTTTCGTGTAACGGTCTTTCGAAAACCTTTATTTTGACAACCAAGTTTATAATTGATTTGTTTTGCTTGACAG

A TGC AAC AAG AGG TCT AGT GGA TTG AGC TAC TCC GAA ATG GAA ACC ACC GAA ACC GTG ATT CTT GGC GTC GGT
s Cys Asn Lys Arg Ser Ser Gly Leu Ser Tyr Ser Glu Met Glu Thr Thr Glu Thr Val Ile Leu Gly Val Gly

CCG GCG AAG ATC CAG TTT GAA GGT GCT GAA ATG AGT GCT GCT TCT GAG GAT GGT GGC TGC AAG TGT GGT GAT AAC
Pro Ala Lys Ile Gln Phe Glu Gly Ala Glu Met Ser Ala Ala Ser Glu Asp Gly Gly Cys Lys Cys Gly Asp Asn
                                                                    Domain 2
TGC ACT TGT GAC CCT TGC AAC TGC AAA TGAAGTGTAACATATAAAAGCTTGAAGCAGAGATATTGAAACCATTATGTTTAATTGTGTGTAT
Cys Thr Cys Asp Pro Cys Asn Cys Lys *

GAGTACACATGTGTTTGGTTTTTTAGAATAATATAATTAATGTGGTCATGGATCTCTCTTCAGCAGATTGAAATATTTGATTTTGTTGAAAGGTTATGAG

AATTGTGGTTGGTTTGTGTTTGGTTTTCTTGTGTAAAGTGTAGCTAAAAGCTTGAATCATAAATCTCTGCATATGATAAATGGAATACATTATTGTGTGT

TATTTGAATGAAAGGATTGTTGGTACAACATATTATGATGTATATTATTAGAGTCGATGATCCAAAGTCTATTTTTTGTGTGAATGAAGATCTAAAAGGT
      ]
AAAAAGTATCTACATTATTTTCTCTCAATTGGCTATAGTGTTTTTTTTAAAGAAGAAAATATGAATTTCTCATTTCATAATCATTCTTTTACCACATGAT
TAGTGACAAGACTATGTAGAGATATGGCCA 3' (+1365)
```

Fig.3. The sequence of *PsMT$_A$* and the predicted amino acid sequence of its product. The sequence of two cDNAs, pPR179 and pPR705, isolated from pea roots were identical to *PsMT$_A$*, minus intron, and the extent of these cDNAs is indicated by square brackets below the gene sequence. The bases are numbered from the translational start site.

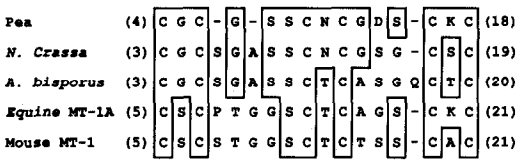
Concurrent with the isolation of *PsMT$_A$*, related cDNAs have been isolated from libraries prepared from poly(A)$^+$ RNA from *Mimulus guttatus* [17]. The pea and *M. guttatus* nucleotide sequences are strongly homologous. Within the coding regions 66% of the bases are identical, and in the regions encoding domains 1 and 2, 74% are identical.

Comparison of the sequences of metal regulatory elements (MREs) of animal MT genes, and examination of the effect of point mutations on their function, have identified the core MRE to be 5′ TGCRCNCX 3′ (where R represents G or A, X represents G or C, and N can be any base but A) [18]. However, in the sixth position where G, C or T give normal regulation, an A gives rise to a high level of basal expression in the absence of metals. The sequence 5′ TGCACACC 3′, flanked by imperfect inverted repeats, occurs between −241 and −248 bases upstream from the start site of

a.



b.



c.

	% amino acid identity to *N. crassa*		% amino acid identity to *Equine* MT-1A	
Pea	59	(80)	53	(60)
N. crassa	100	(100)	41	(41)
A. bisporus	71	(78)	41	(41)
Equine MT-1A	41	(41)	100	(100)
Mouse MT-1	53	(53)	76	(76)
Sea Urchin	59	(59)	29	(29)
Wheat Ec	41	(52)	24	(29)

(in parenthesis, with breaks)

d.

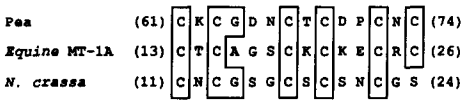


Fig.4. Schematic amino acid comparisons of domains 1 and 2 of the predicted product of $PsMT_A$ and a number of class I and class II MTs. (a) Alignment of domain 1 of the $PsMT_A$ product with regions of 4 class I MTs and 2 class II MTs (sequences obtained from references [6,14]). (b) Alignment of domain 1 with regions of 4 class I MTs allowing for breaks in the sequences. (c) Percentage amino acid identity to *N. crassa* MT and *Equine* MT-1A of the sequences shown in parts (a) and (b). (d) Alignment of domain 2 of the predicted $PsMT_A$ product and class I MTs.

translation in $PsMT_A$ (fig.3). We have no evidence that this element is functional in pea, although there appears to be high basal expression of $PsMT_A$. A sequence ($PsMT_A$, 5′ ATTAAGCATGCAACAATT 3′) with homology (underlined) to part of the 5′-flanking region of the *N. crassa* MT gene occurs between −272 and −290 in $PsMT_A$. Part of this sequence (5′ CATGCAACA 3′) is repeated at −163 to −171. In contrast, there are no sequences in the 5′-flanking region of $PsMT_A$ significantly homologous to the control sequences of *CUP1*, which encodes class II MT from *Saccharomyces cerevisiae* [19].

Pea roots exposed to 50 μM Cd synthesize $(\gamma EC)_nG$ (data not shown). However, the $PsMT_A$ transcript is highly abundant in roots which have not been exposed to high concentrations of trace metals. While $(\gamma EC)_nG$ may detoxify Cd and possibly excesses of other metals in pea roots, the putative $PsMT_A$ product may have a role in essential trace metal metabolism. It is necessary to attempt to purify putative plant MT from pea roots.

Acknowledgements: We warmly thank Professor D. Boulter, in whose laboratory this work was carried out, and Julia Bryden for technical assistance. This work was supported by the Department of Trade and Industry Plant Gene Tool Kit Consortium. N.J.R. is supported by a Royal Society University Research Fellowship.

REFERENCES

[1] Grill, E., Winnaker, E.L. and Zenk, M.H. (1985) Science 230, 674–676.
[2] Jackson, P.J., Unkefer, C.J., Doolen, J.A., Watt, K. and Robinson, N.J. (1987) Proc. Natl. Acad. Sci. USA 84, 6619–6623.
[3] Grill, E., Winnaker, E.L. and Zenk, M.H. (1987) Proc. Natl. Acad. Sci. USA 84, 439–443.
[4] Rauser, W.E. (1980) Nature 287, 563–564.
[5] Robinson, N.J., Ratliff, R.L., Anderson, P.J., Delhaize, E., Berger, J.M. and Jackson, P.J. (1988) Plant Sci. 67, 197–204.
[6] Kagi, J.H.R. and Kojima, Y. (1987) in: Metallothionein II (Kagi, J.H.R. and Kojima, Y. eds) pp.25–62, Birkhauser, Basel.
[7] Evans, I.M., Swinhoe, R., Gatehouse, L.N., Gatehouse, J.A. and Boulter, D. (1988) Mol. Gen. Genet. 214, 153–157.
[8] McMasters, G.K. and Carmichael, G.G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835–4838.
[9] Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
[10] Graham, D.E. (1978) Anal. Biochem. 85, 609–613.
[11] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Hu, N. and Messing, J. (1982) Gene 17, 271–277.
[14] Hofmann, T., Kells, D.I.C. and Lane, B.G. (1984) Can. J. Biochem. Cell. Biol. 62, 908–913.
[15] Kagi, J.H.R. and Schaffer, A. (1988) Biochemistry 27, 8509–8515.
[16] Munger, K., Germann, U.A. and Lerch, K. (1987) in: Metallothionein II (Kagi, J.H.R. and Kojima, Y. eds) pp.393–400, Birkhauser, Basel.
[17] DeMiranda, J.R., Thurman, D.A. and Tomsett, A.B., this issue.
[18] Palmiter, R.D. (1987) in: Metallothionein II (Kagi, J.H.R. and Kojima, Y. eds) pp.63–80, Birkhauser, Basel.
[19] Hamer, D.H. (1986) Annu. Rev. Biochem. 55, 913–951.