Volume 295, number 1,2,3, 171–175 FEBS 10521 December 1991

A plant metallothionein produced in *E. coli*

Peter Kille[1], Dennis R. Winge[2], John L. Harwood[1] and John Kay[1]

[1]*Department of Biochemistry, University of Wales College of Cardiff, Cardiff CF1 1ST, Wales, UK* and [2]*Haematology/Oncology Division, University of Utah School of Medicine, Salt Lake City, UT 84132, USA*

Received 4 October 1991

A metallothionein cDNA was generated from pea (*Pisum sativum* L.) roots, amplified by PCR and inserted into a plasmid for expression in *E. coli*. Purification of the resultant product generated 3 pools of cadmium-containing material after DEAE-cellulose chromatography. The amino acid composition of each was in excellent agreement with that predicted for pea metallothionein. A cadmium content of ~6 g.atoms per mole of protein was estimated. N-terminal sequence analysis revealed that the recombinant molecule had been proteolysed within the extended region linking the 2 cysteine-rich (putative) metal-binding regions. The significance of these findings in terms of the protein folding/targeting of the molecule are considered.

Pea metallothionein; Cloning; PCR; Expression in *E. coli*; Recombinant metallothionein; Extended linker region; Proteolytic nicking; Cysteine-rich region; Metal-binding

1. INTRODUCTION

Metallothioneins (MTs) are widely distributed through species as diverse as fungi and man (for reviews, see [1,2]). Until recently, it was thought that these low mol. wt. metal-binding proteins were not found as such in plants but that phytochelatins (which are not gene-encoded but are synthesised enzymatically) performed equivalent roles [3]. However, 2 recent reports have identified a common mRNA transcript in pea, *Pisum sativum* L. [4] and in the flowering plant, *Mimulus guttatus* [5] which encoded a protein related to metallothionein. These (predicted) proteins differ from archetypal MTs in that (i) their 2 cysteine-rich regions are somewhat shorter (but still align exceptionally well with particular domains of MTs) and (ii) these are separated by a long intervening peptide strand of ~40 residues. In most archetypal MTs, the 2 metal-binding, cysteine-rich domains are connected by a short 'linker' of only ~2 residues [1,2]. Since it has been demonstrated that formation of metal clusters through metal-thiolate liganding prevents the otherwise vulnerable archetypal MT polypeptide(s) against proteolytic attack [6], it might be anticipated that the plant proteins would be susceptible to cleavage within the extended 'linker' region. To date, no direct description of the plant proteins, either from naturally-occurring or recombinant origins, has been forthcoming.

In the present report, we describe the amplification and cloning of the pea MT gene using PCR techniques; expression of the cDNA in *E. coli* permitted isolation of sufficient quantities of this novel MT for initial characterisation.

Correspondence address: J. Kay, Department of Biochemistry, University of Wales College of Cardiff, PO Box 903, Cardiff CF1 1ST, Wales, UK. Fax: (44) (222) 874 116.

2. MATERIALS AND METHODS

Peas (*Pisum sativum* L. cv Feltham First) were grown at 20°C using a 12 h light–dark cycle (650 $\mu E/m^2/s$). After 10 days, the roots were harvested, washed and pooled to give a weight of 5 g of tissue. Total RNA (500 μg) was extracted therefrom using the LiCl Extraction Kit (RPN.1264, Amersham International plc, Bucks., UK) with subsequent repeated precipitations using 3 M sodium acetate to remove residual carbohydrate (D. Grierson, personal communication). Isolation of mRNA (25 μg) from the RNA utilised prepacked oligo(dT)-cellulose columns (Pharmacia, Sweden) and cDNA was synthesised from 5 μg of mRNA using the reagents and enzymes from the cDNA Synthesis System (RPN.1256Y/Z, Amersham). Specific amplification of the pea MT cDNA was achieved using the PCR strategy described previously [7], and sense and antisense primers derived from the internal sequences of the $PsMT_A$ gene [4]. To facilitate cloning, these sequences (Fig. 1a) were extended at their 5' end to encompass a *Cla*I restriction site and an extended Shine-Dalgarno region [8] at the N-terminus, whilst the oligonucleotide encoding the C-terminal region included an *Eco*RI restriction site. Extraction and purification of DNA fragments and authentication of the oligonucleotide sequence was performed as described previously [7].

The pea MT cDNA was inserted into the pPW1 heat-inducible expression vector employing the *Cla*I and *Eco*RI sites as described previously [8]. Expression studies in *E. coli* 1B392 Lon ΔI were performed exactly as described earlier for other MT constructs [8]; cadmium was included at a concentration of 300 μM in the medium 30 min after the incubation temperature had been increased to 42°C. Incubation was then continued at 37°C for 4 h before the cells were harvested and washed. Cell lysis and purification of the recombinant MT from the soluble extract derived therefrom was performed as described earlier [8] with slight variations as documented below.

Amino acid and N-terminal sequence analyses and SDS-PAGE were performed as described previously [8].

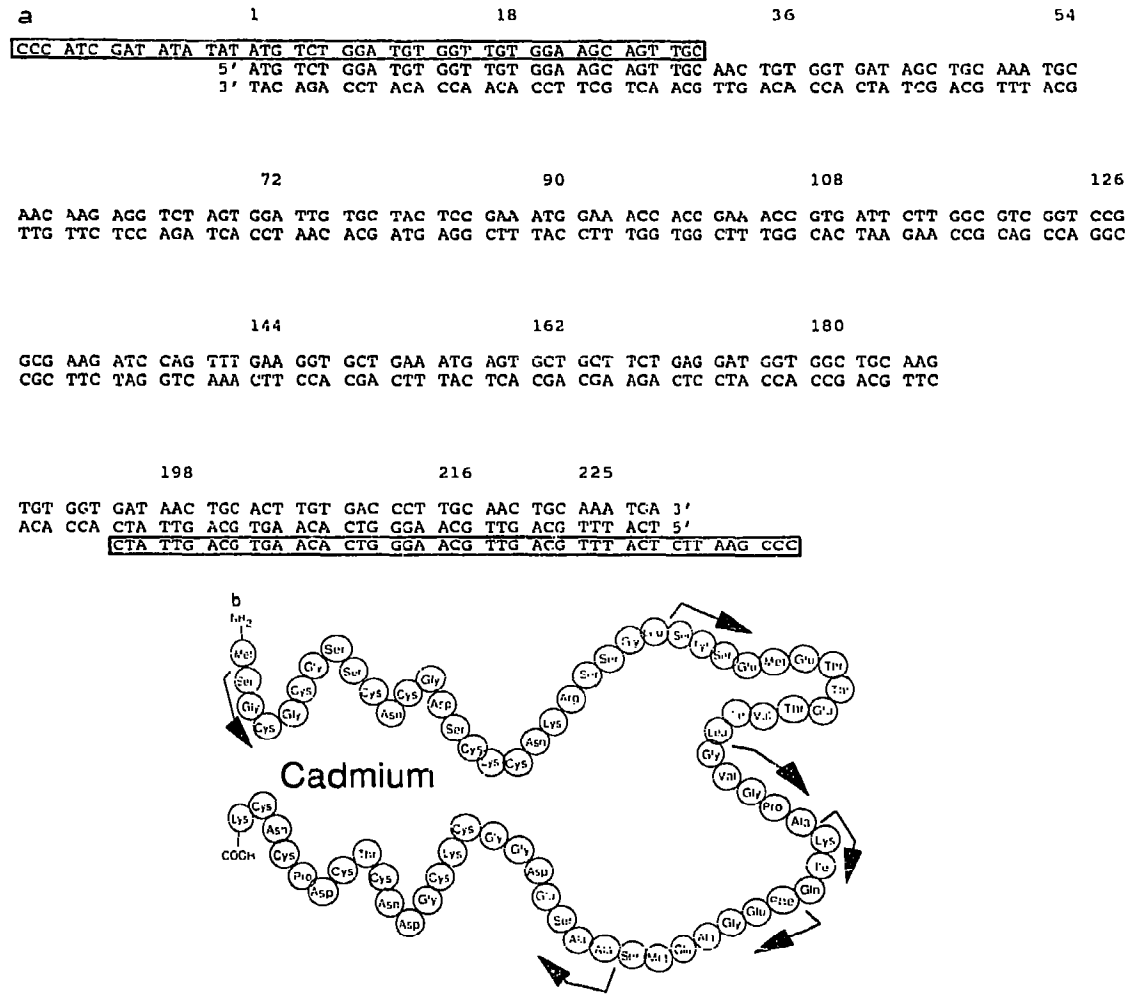

Fig. 1. (a) The nucleotide sequence of the $PsMT_1$ metallothionein gene generated from pea root cDNA by PCR using the oligonucleotide primers highlighted within the boxes. (b) A representation of the pea metallothionein molecule. Sequences identified by N-terminal Edman degradation of the recombinant protein are indicated by →.

3. RESULTS AND DISCUSSION

Analysis of cDNA preparations of 4 positive clones generated as described previously [7] by PCR from pea roots produced a nucleotide sequence in precise agreement (Fig. 1a) with that reported by Evans et al. [4]. Expression of this cDNA in *E. coli* grown in medium containing Cd resulted in an accumulation of intracellular cadmium. As described previously [8], this provides a simple, reliable indicator of the intracellular production of a Cd-binding protein since expression of the wild-type plasmid (without insert) generates no such effect. Eight litres of culture were harvested 4.5 h after heat induction (4 h after Cd addition; [8]) at an E_{610} of 1.2, yielding 42 g (wet wt.) of cells. A cell supernatant was prepared as described previously [8], heated at 65°C for 15 min and centrifuged to remove denatured proteins. The heat-treated supernatant contained 4900 μg of cadmium (representative analysis) of which >90% was retained upon ultrafiltration over a YM-2 membrane. The concentrated extract was applied to a column of Sephadex G-75 (87 × 2.5 cm) equilibrated in 10 mM sodium phosphate buffer, pH 7.4 containing 0.15 M NaCl and 3 mM mercaptoethanol. A single symmetrical peak of Cd-containing material emerged at 72% of the column volume; the peak fractions were pooled, buffer-exchanged into 20 mM Tris-HCl buffer, pH 7.4, containing 3 mM mercaptoethanol and applied to a column of DEAE-cellulose. After washing, desorption was achieved by application of a linear Tris gradient (Fig. 2a).

Analysis of the major Cd-containing fractions by SDS-PAGE (Fig. 2b) indicated the presence of several protein components, all of low molecular weight. The early fractions showed predominantly 1 band with an apparent mol. wt. of approx. 4.5 kDa whilst the later fractions had 2 additional bands migrating as species of ~14 and ~17 kDa, respectively. Since metallothioneins

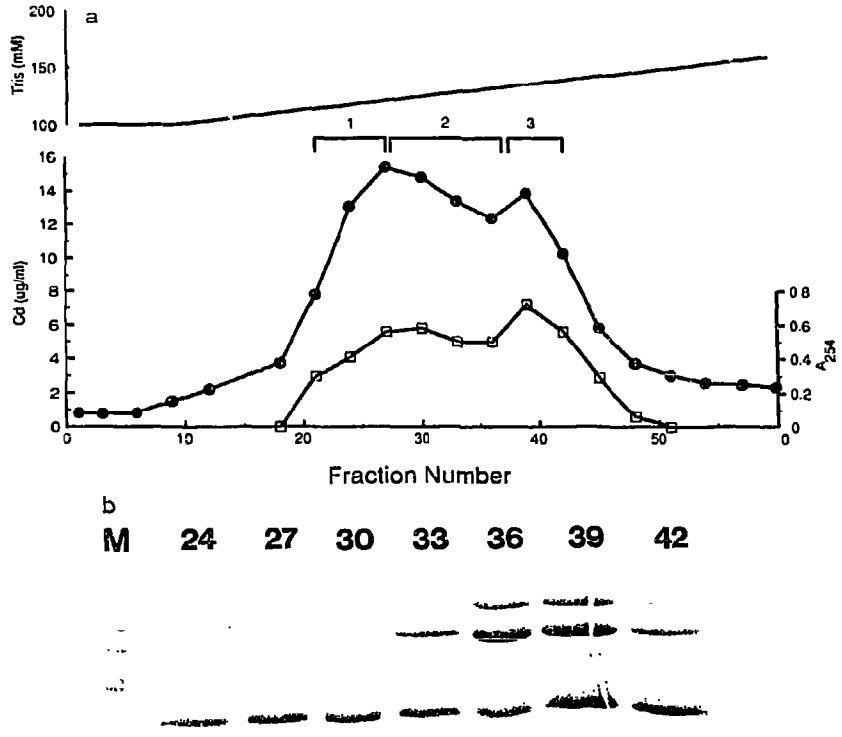

Fig. 2. (a) Ion exchange chromatography of the pooled fractions obtained after gel filtration on Sephadex G-75 of a concentrated extract from *E. coli* expressing a plasmid harbouring the pea MT gene. The cadmium-containing material was applied to a DEAE-cellulose column (14 × 1.6 cm) equilibrated in 20 mM Tris-HCl buffer, pH 7.4, containing 3 mM mercaptoethanol. After extensive washing with this buffer followed by 100 mM Tris-HCl buffer, pH 7.4, containing 3 mM mercaptoethanol, a linear gradient of 100–200 mM Tris-HCl buffer, pH 7.4, containing 3 mM mercaptoethanol (200 ml each) was used for desorption. □= A_{254}; ● = Cd. (b) Aliquots from fractions 24, 27, 30, 33, 36, 39 and 42 from the gradient were analysed by SDS-PAGE on 20% gels. M = markers of mol. wt. 16.9, 14.4, 8.1, 6.2 and 2.5 kDa, respectively. Fractions were pooled as indicated in (a) to generate pools 1, 2 and 3.

are notorious for displaying anomalous molecular weights, fractions were combined as shown (Fig. 2a) and desalted to generate three separate pools. Aliquots of material from all 3 pools were subjected (separately) to gel filtration on a calibrated column of Sephadex G-50 (in 100 mM Tris-HCl, pH 7.4, with 3 mM mercaptoethanol). A single peak of coincident protein and cadmium emerged in all 3 cases at a volume commensurate with an apparent mol. wt. of 14 kDa (not shown). A representative archetypal MT (from human) emerged at an equivalent position (mol. wt. M_r = 13 kDa) from the same column.

Samples of the 3 pools (Fig. 2a) were taken for metal and amino acid analysis and for N-terminal Edman sequencing. The amino acid composition of each pool (Table I) was in excellent agreement with that predicted for pea MT. Metal analyses confirmed the absence of Zn and Cu and indicated a mean content of 5.6, 5.8 and 6.1 g.atoms of cadmium per mole of protein (calculated from the amino acid analysis) for pools 1, 2 and 3, respectively. Representative samples from fractions of pools 1 and 3 were subjected to 5 cycles each of N-terminal sequence analysis.

Several coincident sequences were present throughout the 5 cycles and all were accounted for by the predicted pea MT primary structure (Fig. 1b). The major sequence coincided with that of the authentic N-terminus of pea MT (minus the N-terminal Met residue as observed previously with expression in *E. coli* of recombinant rainbow trout MT; [8]). The additional sequences identified can be assigned to internal locations within the linker region of the pea MT molecule (Fig. 1b) and can be accounted for by multiple proteolytic cleavages within this region, i.e. between the (putative) metal- binding domains. Summation of the total amounts recovered from these 'secondary' sequences generated a value (155 pmol) comparable to that detected for the authentic N-terminal sequence (160 pmol).

To examine these molecular complexities further, aliquots of pools 1 and 2 were exposed (separately) to proteinase K (ratio of 30:1) at 4°C overnight [6]. Closely

comparable results were obtained with both aliquots. Reverse phase FPLC at pH 2 of the digestion mixture caused metal dissociation and enabled resolution of two residual peptides (Fig. 3). Amino acid analysis of these peaks showed the total absence of the aromatic/hydrophobic residues from the linker region of the original molecule (Table I). The compositions can be accounted for by peptides encoding residues 2–21 and 56–75, respectively, from the N- and C-termini of the molecule. Thus, since metal sequestration has been shown previously [6] to protect metal-binding domains in a wide variety of MTs against proteolysis, it would appear that these 2 cysteine-rich regions can indeed serve to chelate metals.

Thus, it would appear that pools 1, 2 and 3 contain only pea MT and, as predicted, the protein was sensitive to proteolytic attack. However, despite such cleavages, the metal-liganding forces within the cysteine-rich regions appear to be sufficiently strong to hold together the 'nicked' molecules thus generated. A single peak was observed on gel filtration using both Sephadex G-50 and G-75 of all 3 pools of material. The extent of proteolytic cleavage within the molecule can account for the various products observed by SDS-PAGE (Fig. 2b). It may well be noteworthy, however, that proteolysis was found to have occurred only within the linker region. No (internal) sequence was detected which began within the cysteine-rich (putative) metal-binding regions of the molecule. If this is the case for the recombinant molecule exposed to the proteolytic machinery of *E. coli*, a similar set of circumstances may occur with

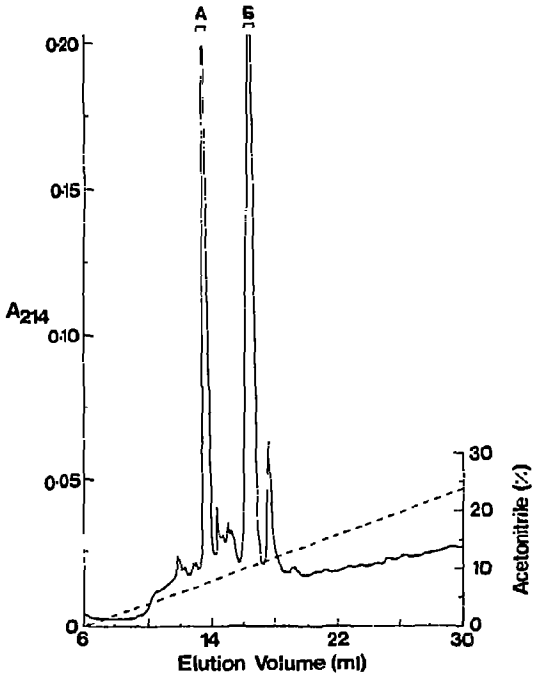

Fig. 3. Reverse phase FPLC at pH 2 on a PepRPC column of pool 1 material (from the DEAE-cellulose column, Fig. 2) following overnight digestion at 4°C with proteinase K. The 2 major peaks eluted by the acetonitrile gradient (at 8 and 11%, respectively) were collected for amino acid analysis. Control samples of undigested pool 1 material and proteinase K were run separately and no distinct major peaks were apparent. Essentially identical results were obtained with material from pool 2.

Table I

Amino acid composition of recombinant pea matallothionein isolated from *E. coli*

	Pool 1	Pool 2	Pool 3	Theoretical whole protein	Peak A	Theoretical residues 2–21	Peak B	Theoretical residues 56–75
Cya	12.6	13.0	11.5	12	NQ	6	NQ	6
Asp	8.0	8.0	8.0	8	3.0	3	4.6	5
Thr	3.6	3.0	3.0	4	0	0	0.9	1
Ser	8.2	8.3	7.5	10	4.3	4	1.3	1
Glu	6.9	7.4	7.5	7	0	0	1.2	1
Pro	2.1	2.4	2.4	2	0	0	0.9	1
Gly	10.7	10.0	8.9	11	3.9	4	3.2	3
Ala	3.8	3.7	3.9	4	0	0	0.2	0
Val	2.1	2.1	2.4	2	0	0	0	0
Met	1.2	1.1	1.1	3	0	0	0	0
Ile	1.7	1.8	2.0	2	0	0	0	0
Leu	1.9	2.3	2.2	2	0	0	0	0
Tyr	0.8	0.7	0.8	1	0	0	0	0
Phe	0.7	0.7	0.8	1	0	0	0	0
Lys	4.7	5.0	4.9	5	1.7	2	2.0	2
Arg	1.0	1.3	1.1	1	0.7	1	0	0

Samples of pools 1, 2 and 3 (Fig. 2a) were hydrolysed for 18, 42 and 66 h in vacuo at 105°C in 6 M HCl following performic acid oxidation. Recoveries for cysteic acid (Cya) were determined by extrapolation to zero time. Samples of peaks A and B from reverse phase FPLC separation of pool 1 material digested with proteinase K (Fig. 3) were hydrolysed similarly but without prior performic acid oxidation. In both cases, cysteic acid was the largest peak present but was not quantified (NQ).

the naturally-occurring gene product(s) exposed to intracellular plant proteinases. This would provide one explanation for the previous difficulties encountered in detecting archetypical MTs in plant cell extracts.

This first demonstration that the protein product expressed from a plant MT gene is capable of metal sequestration thus raises the intriguing question of the purpose of the large linker region of the molecule. Previous experiments in which linker regions of increasing length were engineered into a recombinant mammalian MT [9] demonstrated that (i) up to 12 residues could be introduced without compromising the stability of the molecules expressed in yeast and (ii) increasing the linker region beyond 16 residues reduced the ability of the resultant MT to confer metal resistance on the cells [9]. Since the linker region in the pea molecule is even longer still, with its demonstrated (but expected) susceptibility to proteolytic attack, the function of this segment of polypeptide in folding the metal-binding regions of the molecule into correct juxtaposition or targeting the product into specific sub-cellular locations poses exciting possibilities for further investigation. It does not contain recognisable motifs that have been suggested to direct protein sorting in plants [10]. Aside from these aspects, the 2 cysteine-rich regions, which together are protected by the presence of cadmium against proteolytic trimming by proteinase K, contain 12 thiolate groups. Since all 3 pools of protein contained ~6 g.atoms of Cd/mole, if all of the metal ions are indeed sequestered by thiolate ligands, then of necessity the metal-cysteine connectivities must be different from the ligand pattern observed in archetypal MTs. Structural analysis by NMR spectroscopy will elucidate this further.

Acknowledgements: This work was supported by an SERC/CASE studentship (to P.K.) and by NIH Grant ES 03817 (to D.R.W.). We are most grateful to Prof. Don Grierson (University of Nottingham, UK) for his advice on isolation of RNA from plant tissue, and to Dr. Paul Stephens (Celltech Ltd., UK) for his many valuable contributions to the molecular biological manipulations described herein.

REFERENCES

[1] Kay, J., Cryer, A., Darke, B.M., Kille, P., Lees, W.E., Norey, C.G. and Stark, J.M. (1990) Int. J. Biochem. 23, 1–5.
[2] Kagi, J.H.R. and Schaeffer, A. (1988) Biochemistry 27, 8509–8515.
[3] Rauser, W.E. (1990) Annu. Rev. Biochem. 59, 61–86.
[4] Evans, I.M., Gatehouse, L.N., Gatehouse, J.A., Robinson, N.J. and Gray, R.R.D. (1990) FEBS Lett. 262, 29–32.
[5] de Miranda, J.R., Thomas, M.A., Thurman, D.A. and Tomsett, A.B. (1990) FEBS Lett. 260, 277–280.
[6] Nielson, K.B. and Winge, D.R. (1983) J. Biol. Chem. 258, 13063–13069.
[7] Kille, P., Stephens, P.E. and Kay, J. (1991) Biochim. Biophys. Acta 1089, 407–410.
[8] Kille, P., Stephens, P.E., Cryer, A. and Kay, J. (1989) Biochim. Biophys. Acta. 1048, 178–186.
[9] Rhee, K.-I., Lee, K.S. and Huang, P.C. (1990) Protein Eng. 3, 221–226.
[10] Sebastiani, F.L., Farrell, L.B., Vasquez, M. and Beachy, R.N. (1991) Eur. J. Biochem. 199, 441–450.