Note

Plant Biotechnology, **16** (5), 409-412 (1999)

Isolation of a cDNA Encoding for a Carboxypeptidase, having Leucine Zipper Structure at the N-terminal Region, from the Cultured Shoot Primordia of *Matricaria chamomilla*

Chie KOHCHI[1], Masashi YASUDA[2], and Toshifumi HIRATA[2]*

[1]*Radioisotope Center, Hiroshima University, Higashi-Hiroshima 739-8526, Japan.*
[2]*Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8526, Japan.*
**Corresponding author* E-mail: thirata@sci.hiroshima-u.ac.jp.

Abstract

cDNA encoding serine carboxypeptidase gene of *Matricaria chamomilla* was cloned by reverse transcription-polymerase chain reaction and rapid amplification of cDNA end techniques. The cloned cDNA contained an open reading frame consisting of 501 amino acids, which had two active site motifs of serine carboxypeptidase and high homology with other plant carboxypeptidases. The *M. chamomilla* carboxypeptidase was found to have leucine zipper structure at the N terminal region, suggesting that this enzyme functions at dimer form. Nine *N*-myristoylation sites in the enzyme let us anticipate that this enzyme localizes at a membrane or lipid layer.

It is well known that most plant seeds (sink tissues) store oily substances in discrete organelles termed oil bodies (Bergfeld *et al.*, 1978; Gurr *et al.*, 1974; Murphy and Cummins, 1989). However, we recently found that oil bodies exist even in the multiplying cells (vegetative tissues) of the cultured shoot primordia of *Matricaria chamomilla* (German chamomile), which do not need to accumulate the substances (Hirata *et al.*, 1993; Hirata *et al.*, 1996). On the other hand, it was reported that seed oil bodies contain proteins having protease activities (Kalinski *et al.*, 1990; Qu and Huang, 1990; Vance and Huang, 1987). In the course of the studies on the oil body proteins in the multiplying cells (Izumi *et al.*, 1996), we have investigated the structure of a carboxypeptidase in the cultured shoot primordia of *M. chamomilla* and report herein the deduced sequence of the clone which encodes a polypeptide with active side motifs for serine carboxypeptidase with leucine zipper structure.

According to the reported procedure (Hirata *et al.*, 1993; Takano *et al.*, 1991), shoot primordia of *M. chamomilla* were cultured in test tubes (3 × 20 cm) containing Murashige-Skoog's liquid media (Murashige, 1962) supplemented with 0.02 mg/*l* of α-naphthalene acetic acid and 0.2 mg/*l* of 6-benzylaminopurine at pH 5.7. Cultured shoot primordia of *M. chamomilla* was filtered through 4 layers cheesecloth with aspiration and homogenized in liquid nitrogen using a motor. The total cellular RNA was extracted from the homogenate by the guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987). The RNA was purified from contaminated polysaccharide by precipitation with 2M LiCl.

cDNA cloning was done using reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA end (RACE) method (Loh *et al.*, 1989; Ohara *et al.*, 1989). Until now, several plant carboxypeptidase (CPase) genes have been reported (Baulcombe *et al.*, 1987; Jones *et al.*, 1996; Sorensen *et al.*, 1989; Washio and Ishikawa, 1992). Sequence comparison revealed that those genes are rather conserved among plants and have 60-90% homologies at amino acid level. Thus, we performed an RT-PCR for the partial cloning of CPase of *M. chamomilla* using the primers designed for the conserved regions. The obtained RT-PCR product was composed of 500 bp nucleotides, and the predicted amino acid sequence in the cloned fragment had a high homology with other plant CPases (Baulcombe *et al.*, 1987; Jones *et al.*, 1996; Sorensen *et al.*, 1989; Washio and Ishikawa, 1992). Then, based on the sequence of the cloned 500 bp fragment, the remaining 3' and 5' regions of the gene were also cloned by RACE method. The overall strategy of the cloning and the sequences of the primers were summarized in **Fig. 1** and **Table 1**, respectively. The dT_{12-18} and the 3'RACE-A are primers for the first cDNA synthesis from RNA by reverse transcription. The template cDNA for 5'RACE was prepared by dT_{12-18} primed and dC-

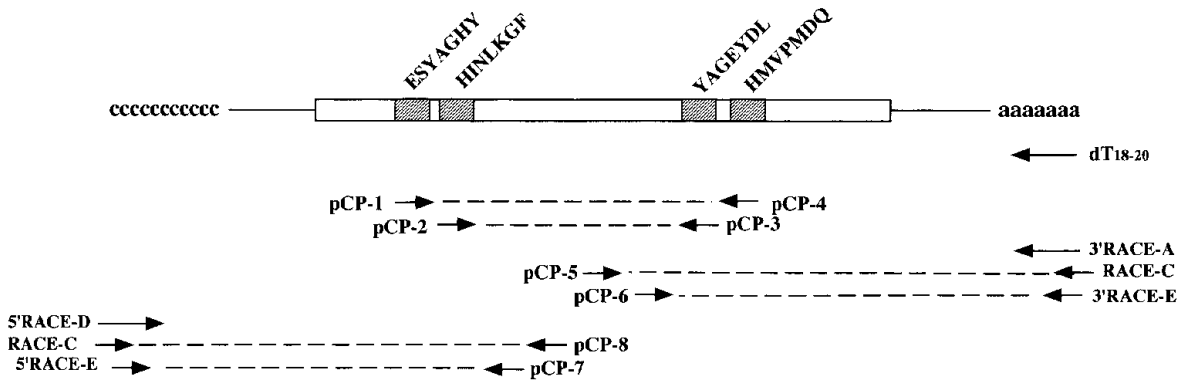

Fig. 1. cDNA Cloning Strategy of *M. chamomilla* CPase (cCPase) using RT-PCR
pCP-1, -2, -3, and -4 primers were designed based on the amino acid sequences E S Y A G H Y (pCP-1), H I N L K G F (pCP-2), Y A G E Y D L (pCP-3), and H M V P M D Q (pCP-4). pCP-5, -6, -7, and -8 were specific primers to the cloned internal region. 3'RACE-A is an oligo-dT primer having Not I adaptor site, and RACE-C and 3'RACE-E have the overlapping adaptor sequences. 5'RACE-D is an oligo-dG primer having Not I adaptor site, and RACE-C and 5'RACE-E have the overlapping adaptor sequences. RACE-C was used commonly in 3'RACE and 5'RACE.

Table 1. Oligo nucleotides used for RT-PCR

Primer	Sequence (5' → 3')
3'RACE-A	TGG AAG AAT TCG CGG CCG CAG TTT TTT TTT TTT TTT TTT
RACE-C	TGG AAG AAT TCG CGG
3'RACE-E	TCG CGG CCG CAG TTT
5'RACE-D	TGG AAG AAT TCG CGG CCG CTT AAG GGG GGG GGG GGG
5'RACE-E	CG CGG CCG CTT A
pCP-1	GAA TCG TAT GCT GGG CAC TA
pCP-2	TCA CAT AAA CTT GAA GGG AT
pCP-3	AA ATC ATA TTC TCC AGC ATA
pCP-4	TG GTC CAT TGG AAC CAT GTG
pCP-5	GTA CAC AGC TAT GCT TGT GG
pCP-6	CCC GAA CTT CTT GAA GAT GG
pCP-7	GGA TCA GTA AGT CCA TTT CC
pCP-8	TAA GCC TGG TAC TGA ATT GC

tailed at the 3' end, followed by second cDNA synthesis using 5'RACE-D as the primer. Molony murine leukemia virus reverse transcriptase (Promega, Madison, USA), terminal deoxy transferase (GIBCO BRL, Rockville, USA), KOD polymerase (TOYOBO, Osaka, Japan), and Pfu turbo DNA polymerase (STRATAGENE, La Jolla, Canada) were used for cDNA preparation and PCR. Each cDNA fragment was cloned in pBluescriptSKII(-).

DNA sequencing of the double-stranded plasmid DNAs were done according to the reported methods (Sanger *et al.*, 1977; Smith *et al.*, 1986) with a DNA sequencing kit for dye terminator cycle sequencing and an Applied Biosystems 377 DNA sequencer using standard protocol of the manufacturer (Perkin Elmer Japan, Applied Biosystems Division, Chiba, Japan). Six independent PCR products were cloned and each sequence was analyzed for the sequence determination.

The obtained 1816 bp-cDNA sequence and the deduced amino acid sequence (GenBank accession # AF141384) are shown in **Fig. 2**. At nucleotide position (ntp) 56-62, a presumable TATA box was found, and the largest open reading frame (ORF), which was composed of 501 amino acids, started at ntp 104. The predicted protein from the largest ORF had two serine carboxypeptidase active site motifs (Stennicke *et al.*, 1996) at amino acid position (aap) 216-223 and aap 455-472, which indicated that the cloned cDNA codes CPase. We termed the cloned cDNA as a *ccp* and the predicted protein as a cCPase. The alignment of the amino acid sequences of cCPase with CPases of *Arabidopsis thaliana* (NCBI accession # 416758) (Jones *et al.*, 1996), *Oryza sativa* (Washio and Ishikawa, 1992) (NCBI accession # 584893), *Hordeum vulgare* (Sorensen *et*

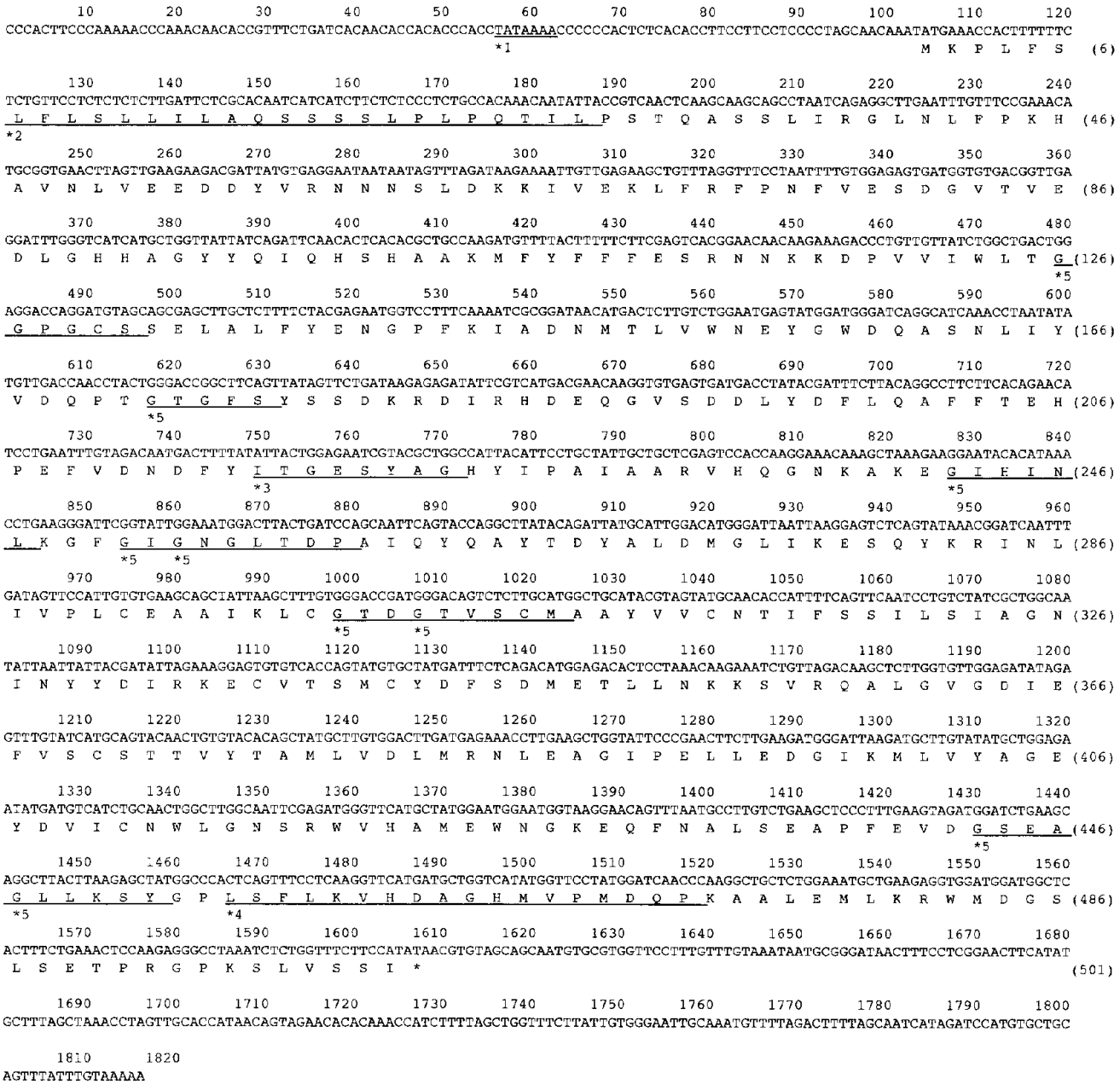

Fig. 2. The Nucleotide Sequence of ccp encoding *M. chamomilla* CPase (cCPase).
The predicted amino acid is shown in below the DNA sequence, assuming that translation begins at the first in-frame methionine of the long open reading frame. Asterisks show motifs as follows; *1 TATA box, *2 Leucine zipper, *3 Serine carboxypeptidase, serine active site, *4 Serine carboxypeptidase, histidine active site, *5 *N*-myristoylation site.

al., 1989) (NCBI accession # 2851577), *Triticum aestivum* (Baulcombe *et al.*, 1987) (NCBI accession # 114874) showed 71%, 76%, 65%, and 66% homologies, respectively.

As a characteristic feature of the cCPase, the leucine zipper structure (Kouzarides and Ziff, 1989; Landschulz *et al.*, 1988; McKnight *et al.*, 1988) was found at the N-terminal region (aap 7-28), suggesting that cCPase exists and/or functions in dimer form. The leucine zipper structure is also found in *A. thaliana* CPase but not in CPase of *O. sativa, H. vulgare* and *T. aestivum*. The leucine zipper structure in *A. thaliana* CPase exists at aap 338-409 but not at N-terminal region (NCBI accession # 416758).

The hydrophobicity plot by the Hopp-Woods method showed no particular feature (data not shown). However, nine possible *N*-myristoylation sites at aap 126-131, 172-177, 242-247, 251-256, 253-258, 299-304, 302-307, 443-448, and 447-452 exist in the amino acid sequence of cCPase, which suggests that the cCPase might localize in a membrane or lipid layer locus. Other plant CPases also have a comparable number of N-myristoy-

lation sites.

Three possible casein kinase II phosphorylation and three protein kinase C phosphorylation sites were found in the cCPase, suggesting that the activity of cCPase might be modified by phosphorylation. Further investigations are needed for clarifying the regulation and function of the cCPase in the shoot primordia of *M. chamomilla*.

Acknowledgments

The present work was in part supported by Grant-in-Aids for Scientific Research No. 09480142 from the Ministry of Education, Science and Culture of Japan, and for the UK-Japan Collaboration Research Project (1997-1998) of Japan Society for the Promotion of Science (JSPS).

References

Baulcombe, D. C., Barker, R. F., Jarvis, M. G., 1987. A gibberellin responsive wheat gene has homology to yeast carboxypeptidase Y. J. Biol. Chem., **262**: 13726-13735.

Bergfeld, R., Hong, Y. N., Kuehnl, T., Schopfer, P., 1978. Formation of oleosomes (storage lipid bodies) during embryogenesis and their breakdown during seedling development in cotyledons of *Sinapis alba* L. Planta, **143**: 297-307.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., **162**: 156-159.

Gurr, M. I., Blades, J., Appleby, R. S., Smith, C. G., Robinson, M. P., Nichols, B. W., 1974. Studies on seed-oil triglycerides. Triglyceride biosynthesis and storage in whole seeds and oil bodies of *Crambe abyssinica*. Eur. J. Biochem., **43**: 281-290.

Hirata, T., Izumi, S., Akita, K., Fukuda, N., Hirashima, T., Taniguchi, K., Nishimori, C., 1993. Formation of oil bodies in cultured shoot primordia of *Matricaria chamomilla*. Plant Tissue Culture Letters, **10**: 289-292.

Hirata, T., Izumi, S., Akita, K., Fukuda, N., Katayama, S., Taniguchi, K., Days, L., Goad, L. J., 1996. Lipid constituents of oil bodies in the cultured shoot primordia of *Matricaria chamomilla*. Phytochemistry, **41**: 1275-1279.

Izumi, S., Takashima, O., Fukuda, N., Hirata, T., 1996. Mr 33K oil body associated protein in culturer shoot primordia of *Matricaria chamomilla*. Phytochemistry, **42**: 309-312.

Jones, C. G., Lycett, G. W., Tucker, G. A., 1996. Protease inhibitor studies and cloning of a serine carboxypeptidase cDNA from germinating seeds of pea (*Pisum sativum* L.). Eur. J. Biochem., **235**: 574-578.

Kalinski, A., Weisemann, J. M., Matthews, B. F., Herman, E. M., 1990. Molecular cloning of a protein associated with soybean seed oil bodies that is similar to thiol proteases of the papain family. J. Biol. Chem., **265**: 13843-13848.

Kouzarides, T., Ziff, E., 1989. Leucine zippers of fos, jun and GCN4 dictate dimerization specificity and thereby control DNA binding. Nature, **340**: 568-571.

Landschulz, W. H., Johnson, P. F., McKnight, S. L., 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science, **240**: 1759-1764.

Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., Davis, M. M., 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. Science, **243**: 217-220.

McKnight, S. L., Landschulz, W. H., Johnson, P. F., 1988. Prediction of a dimerization surface common to a new class of sequence-specific DNA binding proteins. Curr. Top. Microbiol. Immunol., **141**: 186-188.

Murashige, T., 1962. A revisedmedium for rapid growth and bioassays with tabacco tissue cultures. Physiol. Plant, **15**: 473.

Murphy, D. J., Cummins, I., 1989. Seed oil-bodies: isolation, composition and role of oil-body apolipidproteins. Phytochemistry, **28**: 2063-2069.

Ohara, O., Dorit, R. L., Gilbert, W., 1989. One-sided polymerase chain reaction: the amplification of cDNA. Proc. Natl. Acad. Sci. U S A, **86**: 5673-5677.

Qu, R. D., Huang, A. H., 1990. Oleosin KD 18 on the surface of oil bodies in maize. Genomic and cDNA sequences and the deduced protein structure. J. Biol. Chem., **265**: 2238-2243.

Sanger, F., Nicklen, S., Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U S A, **74**: 5463-5467.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B., Hood, L. E., 1986. Fluorescence detection in automated DNA sequence analysis. Nature, **321**: 674-679.

Sorensen, S. B., Svendsen, I., Breddam, K., 1989. Primary structure of carboxypeptidase III from malted barley. Carlsberg Res. Commun., **54**: 193-202.

Stennicke, H. R., Mortensen, U. H., Breddam, K., 1996. Studies on the hydrolytic properties of (serine) carboxypeptidase Y. Biochemistry, **35**: 7131-7141.

Takano, H., Hirano, M., Taniguchi, K., Tanaka, R., Kondo, K., 1991. Rapid clonal propagation of *Matricaria chamomilla* by tissue-cultured shoot primordia. Japan. J. Breed, **41**: 421-426.

Vance, V. B., Huang, A. H., 1987. The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. J. Biol. Chem., **262**: 11275-11279.

Washio, K., Ishikawa, K., 1992. Structure and expression during the germination of rice seeds of the gene for a carboxypeptidase. Plant Mol. Biol., **19**: 631-640.