Isolation and structure determination of a cDNA encoding for a thiol protease from the cultured shoot primordia of *Matricaria chamomilla*

Chie KOHCHI¹, Masashi YASUDA², Toshifumi HIRATA^{2*} and Lionel John GOAD³

¹Radioisotope Center, Hiroshima University, Higashi-Hiroshima 739-8526, Japan ²Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8526, Japan ³School of Biological Sciences, Life Sciences Building, University of Liverpool, Liverpool L69 7ZB, UK *Corresponding author E-mail address: thirata@sci.hiroshima-u.ac.jp

Received 4 February 2000; accepted 24 April 2000

Abstract

A cDNA encoding for a thiol protease was isolated from *Matricaria chamomilla*. The cDNA contained an open reading frame consisting of 501 amino acids, which had three active site motifs of thiol protease.

Most living organisms have developed defense reaction to protect themselves against exogenous stimulus. As one of the defense reactions, intracellular enzymes are secreted from the plant cells in response to an exogenous stimulus (Chibbar et al., 1984; Izumi et al., 1995). We have reported that various monoterpenoids, especially geraniol, exhibit a potent activity for the induction of apoptosis-like cell death as a defense reaction (Izumi et al., 1999) in the cultured shoot primordia of M. chamomilla (German camomile). Recently we found that several proteases are secreted from the cultured shoot primordia of M. chamomilla when geraniol was administered to the cultures (Izumi et al., 1996; and unpublished data). In continuation of the structure determination of proteins concerning to the defense reaction in higher plants (Kohchi et al., 1999), we have now investigated the protein with protease activity in the cultured shoot primordia of M. chamomilla and report here the primary structure of a thiol protease.

Shoot primordia of *M. chamomilla* were cultured as described elsewhere (Hirata *et al.*, 1993; Takano *et al.*, 1991). Total cellular RNA was isolated by guanidine thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987). Contaminated polysaccharide was removed from RNA by precipitation with 2M LiCl.

cDNA cloning was performed using reverse transcription-polymerase chain reaction (RT-PCR) and its modified method, i.e., the rapid amplification of cDNA end (RACE) method (Loh *et al.*, 1989; Ohara *et al.*, 1989). The sequences of PCR primers other than Oligo dT_{12-18} (Pharmacia Biotech, Tokyo, Japan) and the schematic positions of the primers on a supposed mRNA for M. chamomilla thiol protease are summarized in Table 1 and Fig. 1, respectively. Moloney murine leukemia virus reverse transcriptase (Promega, Madison, USA), terminal deoxy transferase (GIBCO BRL, Rockville), and KOD polymerase (TOYOBO, Osaka, Japan) were used for cDNA preparation and PCR. Each PCR product was cloned in pBluescriptSKII(-). DNA sequencing of double stranded plasmid DNAs was conducted according to the method reported previously (Sanger et al., 1977; Smith et al., 1986) and the standard manufacturer's protocol with a DNA sequencing kit for dye terminator cycle sequencing using an Applied Biosystems 377 DNA sequencer (Perkin Elmer Japan, Applied Biosystems Division, Chiba, Japan).

Recently, we isolated a protease from cultured shoot primordia of M. chamomilla as a defenserelated protein against chemical stress (unpublished data). The partial amino acid sequence of the protease (PEAAHAINTG) was found to have homology with a region (IEAAHAIATG) of Glycine max thiol protease (TPase) (Kalinski et al., 1990). Thus, in an attempt to clone TPase gene of M. chamomilla, we first performed RT-PCR using degenerated primers corresponding to the following amino acid sequence, PEAAHAINTG (pTP-1), QELVDCV (pTP-2), WIAKNS (pTP-3), and WGEDWG (pTP-4). The latter three amino acid sequences are those of G. max TPase, and they position downstream of IEAAHAIATG (Table 1, Fig. 1). As illustrated in Fig. 1, cDNA reverse transcribed with Oligo dT12-18 was applied to 1st

Primer	Sequence $(5' \rightarrow 3')$
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'TACE-D	TGG AAG AAT TCG CGG CCG CTT AAG GGG GGG GGG GGG
5'RACE-E	CG CGG CCG CTT A
pT P -1	CCI GA(AG) GCI GCI CA(CT) GCI AT(ACT) AA(CT) AC(ACGT) GG
pT P -2	CAA GAG CTC GT(AT) GAC TGT G
pTP-3	CA (TC)GA GTT CTT GA(TC) GAT CCA
pTP-4	CC CCA (AG)T(CT) T(GT)C TCC CCA (TC)GA
pTP-5	GGG TTC GGC TTA TGA TTT CC
pTP-6	ACA CCG GGG GTG TAT ATA AC
pTP-7	CGC ACC CAT AGT CAT AAG TG
pTP-8	CGT GTC CAT GTT TCC ACC AT
pTP-9	CTC TTC ATT ACT CAA GTC AGC

 Table 1.
 Oligo nucleotides used for RT - PCR



Fig. 1. The schematic positions of primers used for PCR on a supposed mRNA of *M. chamomilla* thiol protease.

A supposed mRNA and its ORF (dotted line square) for *M. chamomilla* thiol protease (A), and a mRNA and its ORF (solid line square) for *G. max* thiol protease (B) are shown together with their partial amino acid sequences (gray squares). Arrows show primers, their directions, and schematic positions corresponding to mRNAs above.

PCR using pTP-1 and pTP-4 as primers. An aliquot of the 1st PCR product was subjected to a second nested PCR using pTP-2 and pTP-3. The nested PCR gave a 410-bp DNA fragment. Sequence analysis revealed that the DNA fragment is a part of an open reading frame (ORF), in which an eukaryotic thiol protease histidine active site is contained. Thus, we named the cDNA as *ctp* regarding as a cDNA encoding for one of the thiol proteases of *M. chamomilla*.

To clone the whole ORF of ctp, we subsequently tried 3' and 5' RACE-PCR. For 3' RACE-PCR, two sense primers (pTP-5 and pTP-6) were designed based on the determined ctp sequence, and three antisense primers (3'RACE-A, RACE-C, and 3'RACE-E) having Not I adapter site were prepared (**Table 1**). cDNA reverse transcribed with 3'RACE-A was applied to 1st PCR using pTP-5 and RACE-C as primers. An aliquot of the 1st PCR product was subjected to a second nested PCR using pTP-6 and 3'RACE-E (**Fig. 1**). The nested PCR gave DNA fragments which have polyA stretch at the 3'-end and are overlap with ctp at 5'-end.

For 5' RACE-PCR, two antisense primers (pTP-7 and pTP-8) were designed from the determined *ctp* sequence, and three sense primers (5'RACE-D, RACE-C, and 5'RACE-E) having Not I adapter site were prepared (**Table 1**). Oligo dT_{12-18} primed single stranded cDNA was dC tailed at the 3' end and then the double stranded cDNA was synthesized using oligo dG/Not I adaptor primer (5'RACE-D). First PCR was performed using RACE-C and pTP-8 as primers. An aliquot of the 1st PCR product was subjected to a 2nd nested PCR using

5'RACE-E and pTP-7. After confirming that the nested PCR product has a cDNA overlapping with *ctp*, pTP-9 primer was newly prepared. To determine the sequence of 5' non coding region of *ctp* more longer, 5' RACE-PCR was re-performed starting with cDNA reverse transcribed with pTP-8 followed by 5'RACE-PCR using pairs of primers, RACE-C / pTP-7 and 5'RACE-E / pTP-9 (Fig. 1).

Sequences of PCR products were analyzed for at least 6 independent clones for each region and consensus sequences were determined, respectively. **Fig. 2** shows the combined 1732 bp-entire sequence of *ctp* and predicted amino acid sequence (GenBank accession # AF182079).

The largest ORF, which is composed of 501 amino acids, starts at nucleotide position (ntp) 44. The predicted protein from the ORF has three eukaryotic thiol protease motifs, i.e., cysteine active site at amino acid position (aap) 161–172, histidine active site at aap 305–315, and asparagine active

site at aap 322-341. We designated the predicted protein as cTPase. The cTPase may not be identical to the protease which we have previously purified from oil body of *M. chamomilla*, because PEAAH-AINTG sequence is not found in the deduced amino acid sequence. Instead of PEAAHAINTG, a related amino acid sequence, IESANAIATG is found at aap 176-185 in cTPase. Presumably, this region appears to cross react with the degenerated PCR primer, pTP-1, so that *ctp* was amplified by RT-PCR.

Homology search in the protein database indicated that cTPase has approximately 40% homology with TPase of Lycopersicon esculentum (NCBI # AJ003137) (Lers et al., 1998), Phaseolus vulgaris (NCBI # Z99954), Solanum tuberosum (NCBI # AJ245924), and Glycine max (NCBI # J05560) (Kalinski et al., 1990) as shown in Fig. 3. These plant TPases have a high number of N-myristoylation sites in their molecules. cTPase also has 14 possible N-myristoylation sites. This suggests that the plant TPases, including cTPase, might localize

	10		20		30)		40		50			60	~ ~ ~		70			80			90_		10	00		1	10		12	0
AAAAGA	GCACCF	ATGC	CAAGA	TTCA	TTAGC	TTC	ACTA	AGCTA	AAAT M	A I	S	AAA: N	S	M	GAT I	T	ATTC I L	FCAT: I	F	CTT/ L	ACC T	TAT Y	GTTT V S	CATF Y	ACTO S	IATC I	SATC.	AACI	K	ACAC: T L	r (26)
																					_						_				_
тестая	130 TGAGTI	TTCT.	140 ATACT	TGAA	150 GGTCA) AGAI	המאמ	160 IGATA	TTCT	170 ATCAA	GTGC	:AAA	180 AGT	TAG	TGA	190 ССТА	TTTG	2 3888	00 47G	GAA		10 ፐፐር	CATC	22 GAA2	20 AAA(сата	2	30 ACA'	TGAA	24(GAAG	D A.
PS	EF	s	ΙL	Е	GQ	Е	N	DI	L	s s	A	ĸ	v	s	D	L	FG	ĸ	W	ĸ	E	L	H G	K	T	Y	Q	H	E	ΕĒ	(66)
	250		260		270)		280		290			300			310		3	20		Э	30		34	10		3	50		36	0
GGAGAA	CCTTAC	GCTT	GAGAA	TTTC	AAGAA	AAG	IGTA	AAGT	TTGT	AATGG	AGAA	AAA	CTC.	AGA	GAG	GAAG	TCTG	AGCT	AGA	TCAT	ACT	GTG	GGAT	TGA	TAI	AGTI	TGC	TGA	CTTG	AGTA	A
EN	LR	L .	e n	F	КК	s	v	KF	v	ME	ĸ	N	s	Е	R	ĸ	SE	L	Ð	В	т	v	GL	N	ĸ	F	A	D	L	S N	(106)
	370		380		390)		400	~~ ~	410		* ***	420	~~ ~~	~~~	430		4	40		4	50		46	50		4	70		48	D
TGAAGA E E	GTTTAF F K	AGAG	ATGTA MY	M	TCAAA S K	AGT.	raag K	GGGGT G S	CTAG R	AAGTA S N	ATGA	L TAL	AAA K	GATO	GGG	GGG1	GTGAL V K	AAAGO R	SAA N	CATC	SAGT	GTA V	AGTT S S	CGAG	igat T	CTTG	TGA	TGC/	P	ACTTC T S	2 (146)
		5.				•					~	-		••	Ű	· · ·	• •				- -		0 0			Ū	Ű.		• ·		
ተምምርር እ	490 	acac	500	AGTA	510 הדדמר) 'acci	ል ልጥ ር	520	<u>አ</u> ግሮ አ	530 2000	ልርጥር	maa	540	TTG	ዋጥርራ	550 260 a	m m cr m cr	5) המכדע	50 200	TCCI	5 100 a	70 አምሮ	2333	58 התוכר	30 "5 5 7	MAC	5 ידממי	90 2000	CACC	600 נמדימים	
L D	WR	D	KG	v	V T	P	M	KD	Q	GΟ	C	G	S	c	W	A	FS	V	s	G	s		(S	A	N	à		A	T C	D	(186)
	~ ~ ~		6.0.0					<i></i>	*1	CE 0						670			-			~ ~	(pTP	-1)				10		-	
CCTCAT	TAGACI	CTCA	620 GAGCA	AGAG	630 CTTGT	CGA	TGT	GACA	CTTA	050 TGACT	ATGO	GTG	CGA	TGG	TGG	670 AAAC	ATGG	ACACO	30 GC	ттат	o 'AGA'	90 TGG	מידיא	70 קמידי	JU IGAZ	ATGG	. ז הפסידו	IU GCTr	rgam'	/ZU TCTG/	
LI	R L	SI	E Q	8 1	L V	D	C	1 П	Y	DΥ	G	С	D	G	G	N	M D	т	A	Y	R	W	ΙI	ĸ	N	G	G	L	D	S E	(226)
	720		740	(pTP-	2)			760		770			700			200						10						20			
AGACGA	7 3U TTATCC	ATAC	ACCAG	ттсти	AATGG		CGAT	GGTA	AATG	TGACA	ааас		GTC	AGC	AAA	790 GTCA	GTTG		JU GCT	TGAT	AGC	TAT	GTGG	82 AAGT	:U 'AG#	ATC	GAA	3U TGAJ	AGAT	640 GCAG	, ,
D D	ΥP	Y	г S	SI	N G	R	D	GK	С	DK	Т	ĸ	s	A	K	s	vv	S	L	D	s	Y	V E	v	E	s	N	E	DJ	A V	(266)
	0=0		000		070			000		000						010					0	20						F 0		00	_
TTTATG	TGCTGI	GGCA	ACTAC	CCCT	GTTAC	TAT	IGGT	ATCG	TGGG	090 TTCGG	CTTA	TGA	900 TTTC	CCAC	GCT	ATAC	ACCG	GGGG	CGT.	атат	'AAC	30 GGA	CAAT	94 GCTC	IAAC	таа	ACC:	JU ATAT	FGAT:	901 ATAG	, ,
L C	A V	A !	гт	P	νт	I	G	ΙV	G	S A	Y	D	F	Q	L	Y	T G	G	v	Y	N	G	çс	S	S	ĸ	P	Y	Dj	<u>[D</u>	(306)
	970		980		990		1	000		1010		1	020			1030		104	10		10	50		106	0		10	70	,	*2 108()
CCATGO	AGTGCI	AATCO	STTGG	TTAT	GGTTC	ACA	AGAT	GGCA	AGGA	CTACT	GGAI	CGT	CAAG	GAAI	FTC	ATGG	GGCA	TTAT	TG	GGGA	CTA	GAA	GGCT	ACAT	TCI	гаат	GGA	AAG	4AAC/	ACTG?	A .
<u>H A</u>	VL	IV	/ G	ΥĢ	<u>s</u> s	Q	D	GK	D	<u>Y</u> W	1 /_T	V D 2	K	N	S	H	3 I	<u> </u>	W	G	LE	5 6	<u>; y</u>	I	L	M	E	R	N T	Ð	(346)
	1090		1100		1110	•	1	120		1130	(pr.	1	140			1150	(prP-	4) 116	50		11	70		118	10		119	90		120()
TATTAA	AAACGG	GGTT	IGTGG	AATG	TACCI	TGAG	GCCG	GTGT.	ACCC	TATCA	CTGC	GGC	GCC	AAC!	ACCO	CCCT	GGTC	CACCO	CC.	ACCA	CCA	GCC	CCGC	CTTC	TCO	CACC	ACA'	FCCF	ACCTO	CACC	2
IK	NG	v	G	M	ΥL	Е	Р	VY	Р	IT	A	A	<u>P</u> * 4	Т	P	P	G P	P	Р	P	P /	A 1	<u> </u>	S	P	P	H	P	<u>P</u> P	, <u>p</u>	(386)
	1210	:	1220		1230	•	1	240		1250		1	260		:	1270		128	30		12	90		130	0		13	10		1320)
CCCAAC	GCCACC	AGCA	CCTAG	CAAG	FGTGG	TGA	CTTC	CATT	ATTG	TGCGG	CTGA	CCA	GAC	ATGI	FTG:	FTGC	ATAT	TGAG	TT	CTAI	AAC	TAT	TGTT:	FGAT	CTF	ATGG	TTG	ITGI	IGGT?	PATTO	
<u> </u>	PP	A	<u> </u>	ĸ	5 G	D	P.	HY	С	AA	D	Q	T	C	C	C	L F	Б	Ŀ.	Y	N 1	r (1	Y	G	Ċ	C	GY	S	(426)
	1330		1340		1350		1	360		1370		1	380			1390		140	00		14	10		142	0		14	30		1440)
CGATGC	CGTGTG	TTGC	AAGAA	TAGTO	GCTGC	TTGI	CTGT	CCTA	GCGA	TTACC	CCAI T	TTG	TGA	CGTI	ICA)	AGCT	GGAT <i>i</i> G V	ACTGO	CTA'	TAAG	N	CTC	GCAA	AGAC	GTI: F	rcgg	GGT	3CCA	AGCA	AAGA#	4661
5 6	• •	· ·		0 1		C	C	1 0	5	• •	-	C	D	•	×		5 I	C	-					1		0	×	r	^ 1	·	(400)
	1450		1460		1470		1	480		1490		1	500			1510		152	20		15	30		154	0		155	50		1560)
ACGACA R O	ACTAGC	AAAG(K 1	JACAA H K	GATG(M 1	JCATC P W	GGA	JAAG K	ATAG.	AAGA E	AACAA T T	TCAA K	IGGA E	GGA(E	STTC F	CAC 0	P	CTTGO	GTGC: W	SAA' N	TAGG R	IAACI N	UCG' P	TTTG(F A	LAGC A	AGC: A	GGC: A	1'I'A2 *	AGAP	ICAAO	JACAI	(501)
~~ ¥						-			-			~	~	-	×	-															,
# A C C C A	1570		1580	~~ ~ ~~	1590	2000	1 • • • • •	.600	ካ አጥ/?	1610 TCTAT	C TTC TT	1	620 GNR	n a ma	ג ידייי	1630	ጠ መጥ ርጉ ን	164	10 - A TH	« ۱۳۰۵	16	50	7 7711 2 /	166	0	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	16	70		168()
INGGGA	TALLL	TAGC.	IGATC	CCTT	aceri	RGGG	3GTT	ATGT	nw10	LOTAL	9101	TTA	GMA.	THT	114(~~~~A	+1104	NUW L(Ar	GIMA	icer.	nnc)	CGTA	JGT H	en n n	CIT	GAAG	ACE	rear	3TCH0	,
	1690	:	1700		1710		1	720		1730		1	740																		
GTGTGA	AATTTG	ACGA	IGAAA	AAAA	AAAAA	AAA	AAAA	AAAA	аааа	ААААА	AA																				

Fig. 2. The nucleotide sequence of *ctp* encoding *M. chamomilla* TPase (cTPase).

The predicted amino acid sequence is shown below the DNA sequence. Asterisks show motifs as follows; *1 Eukaryotic thiol protease cysteine active site, *2 Eukaryotic thiol protease histidine active site, *3 Eukaryotic thiol protease asparagine active site, *4 proline rich region. Shadows represent the possible amino acid loci cross-reacting with pTP-1, 2, 3, and 4.

M.c.	1	MATSNSMITILIFLTYVSYSISTKTLPSEFSILEGQ-ENDILSSAKVSDLFGKWKELHGKTYQHEEEENLRLENFKKSVKFVMEKNS	86
L.e.	1	${\tt MAAHSSTLTISILLMLI-FSTLSSASDMSIISYDETHIH-RRTDDEVSALYESWLIEHGKSYNALGEKDKRFQIFKDNLRYIDEQNS$	85
P.v.	1	MLLFALFA-LSSALDMSIISYDNAHQD-KATWRTDBEVNSLYEEWLVKHGKLYNALGEKDKRFQIFKDNLRFIDQQNA	76
s.t.	1	MAAHSSTLTISLLLMLI-FSTLSSASDMSIISYDETHIH-HRSDDEVSALYESWLIEHGKSYNALGEKDKRFQIFKDNLKYIDEQNS	85
G.m.	1	MG-FLVLLLFSLLGLSSSSSISTHRSILDLDLTKFTTQKQVSSLFQLWKSEHGRVYHNHEEEAKRLEIFKNNSNYIRDMNA	80
		·····*···*···*···*···*···*···*···*···*	
M.c.	87	ERKSELDHTVGINKFADI SNEEFKEMYMSKVKGS-RSNELKMGGVKRNMSVSSRTCDAPTSLDWRDKGVVTPMKDOGOCGSCWAFSVSGS	175
L.e.	86	VPNOSYKIGITKFADLTNEEYRSIYLGTKSSGDRKKLSKNKSDRYLPKVGDSLPESIDWREKGVLVGVKDOGSCGSCWAFSAVAA	170
P.v.	77	ENR-TYKLGLNRFADLTNEEYRARYLGTKIDPNRRLGRTPSNRYAPRVGETLPDSVDWRKEGAVVPVKDOASCGSCWAFSAIGA	159
s.t.	86	VPNOSYK LGLTKFADI TNEEYRSTYLGTKSSGDRRKLSKNKSDRYLPKVGDSLPESVDWRDKGVL/GVKDGSCGSCWAFSAVAA	170
G m	81	NEKSPHSHELGLNKFADTTPOEFSKKYLOAPRDVSOOTKMANKKMKKEOYSCDHP-PASWDWRKKGVTTOVKYOGGCGRWAFSATGA	167
	•••	······································	
		(pTP-1) (pTP-2) TP motif (cysteine)	
MC	176	TES ANA TATION LEL SEDERATION - VDVGCDGGNMDTA VEWT IKNGGLDSEDDVPYTSSNGEDGKCDKTKSAKSVVSLDSVVEVESNED	264
T. e	171	MESTNA LYTCHI, ISI, SECEL WOODRSYNEGCOGGI MDY AFREVI KNGGT DTREDY PYKENGCOOYBKNAK-WKI DSYRDYPVNNE	257
P.v	160	VIRGENKETVERGEDE VIRGEN VOC DE VIRGEN VOC VIRGEN VERKETEN VOC VIRGEN VOC VI VIRGEN VIRGEN VOC VIRGEN V	246
s.+	171	MSCINALIVICALLISI.SEOELUDCORSVIRGCOGGLMOVAEREVILNINGCIDTERDVPYKRENDVCDOVRKNAK-VVKIDSVEDVPVNNE	257
G.m.	168	TRAAHA LATGDI VSI.SEORTADCYFESEGSYNGWOYO-SFEWYLEHGGI ATDDDYPYRAKEGRCKAN-KIODKYTIDGYETLIMSDESTE	255
	100	***************************************	
		(pTP-3) (pTP-4)	
M.c.	265	AVLCAVATT-PVTIGIVGSAYDFQLYTCGVYNGQ-CSSKPYDIDHAVLIVGYGSQDCKDYWIVKNSWCTYWCLEGYILMERNTDIKN-	349
L.e.	258	KALQKAVAHQPVSIALEAGGRDFQHYKSGIFTGK-CGTAVDHGVVIAGYGTENGMDYWIVRNSWGANWGENGYLRVQRNVASSS-	340
P.v.	247	${\tt LALKKAVANQPVSVAVEGGGR-=EFQLYSSGVFTGR-CGTALDHGVVAVGYGTDNGHDFWIVRNSWGADWGEEGYIRLERNLGNSRS}$	330
s.t.	258	KALQKAVAHQPVSIAIBAGGRDLQHYKSGIFTGK-CGTAVDHGVVAAGYGSENGMDYWIVRNSWGAKWGEKGYLRVQRNVASSS-	340
G.m.	256	SETFQAFLSAILEQPISVSIDAKDFHLYTGGIYDGENCTS-PYGINHFVLLVGYGSADGVDYWIAKNSWGEDWGEDGYIWIQRNTGNLL-	343
		······ ·······························	
		TP motif (higtidine) TP motif (asparagine)	
Ма	350		438
п.с.	3/1	Groci A TERSYDIK WARDNOK PARSON KOW DOWN DOWN THE THE OTOM THE OTOM THE AND THE THE THE OTOM THE AND T	419
D. 17	241	CHOREN THE SAME AND	400
£.v.	241		410
5.C.	241		379
G.m.	344	Geographical Filesetti Sarvioner (Dest L	373
		· · · · · · · · · · · · · · · · · · ·	
M.c.	439	PSDYPICDVQAGYCYKNSAKTFGVPAKKRQLAKHKMPWEKIEETIKEEFQPLAWNRNPFAAAA	501
L.e.	420	PHDYPICNVRQGTCSMSKGNPLGVKAMKRILAQPPHDYPICNVRQGTCSMSKKSSS	466
P.v.	410	PHDYPICNTYAGTCLRSKNNPFGVKALRRTPAKPHGAFAGN-KVSNA	455
s.t.	420	PHDYPVCNVRQGTCSMSKGNPLGVKAMKRILAQPPHDYPVCNVRQGTCSMSKKSSS	466
G.m.	379		379

Fig. 3. Comparison of amino acid sequences of thiol protease among several plants.

M.c.: *Matricaria chamomilla*, L.e.: *Lycopersicon esculentum* (NCBI # AJ003137) (Lers et al., 1998), P.v.: *Phaseolus vulgaris* (NCBI # Z99954), S.t.: *Solanum tuberosum* (NCBI # AJ245924), G.m.: *Glycine max* (NCBI # J05560) (Kalinski *et al.*, 1990). Asterisks denote identical amino acid residues through all thiol proteases and dots denote common amino acid residues to three of five thiol proteases. Dashes represent gaps. The thiol protease active sites are shown below the amino acid sequences. Shadows represent the possible amino acid loci cross-reacting with pTP-1, 2, 3, and 4.

in membrane or lipid layer locus. Although cTPase would be a different protease from that we previously isolated from *M. chamomilla* (Izumi *et al.*, 1996), it could be one of the membrane- or oil body-associated proteases. As a characteristic feature of cTPase, a proline rich region is found at aap 366-392, but the role of the proline stretch is not clear. A number of possible target sites for casein kinase II, protein kinase C, and tyrosine kinase are found in cTPase, suggesting that the activity of cTPase might be modified by phosphorylation.

Further investigations are needed for clarifying the function of cTPase in the shoot primordia of *M. chamomilla*.

Acknowledgements

The present work was in part supported by Grantin-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

Chibbar, R. N., Cella, D., Albani, D., Huystee, R. B. V., 1984. The growth and peroxidase synthesis of two carrot cell lines. J. Exp. Botany, **35**: 1846-1852.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., **162**: 156159.

- 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.
- Izumi, S., Takashima, O., Fukuda, N., Hirata, T., 1996. Mr 33K oil body associated protein in cultured shoot primordia of *Matricaria chamomilla*. Phytochemistry, 42: 309-312.
- Izumi, S., Takashima, O., Hirata, T., 1999. Geraniol is a potent inducer of apoptosis-like cell death in the cultured shoot primordia of Matricaria chamomilla. Biochem. Biophys. Res. Commun., 259: 519-22.
- Izumi, S., Yamamoto, Y., Hirata, T., 1995. Secretion of an esterase from the cultured suspention cells of *Marchantia polymorpha*. Phytochemistry, 38: 831-833.
- 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.
- Kohchi, C., Yasuda, M., Hirata, T., 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*. Plant Biotech., **16**: 409-412.
- Lers, A., burd, S., Sonego, L., Khalchitski, A., Lomaniec, E., 1998. Nucleotide sequence of a full-length C14 cDNA clone encoding for cysteine protease CYP1 from tomato. Plant Physiol., 116: 1193-1193.
- 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. Sci., 243: 217-220
- 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.
- 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.
- 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.