## Isolation and characterization of cDNA clones for carrot extensin and a proline-rich 33-kDa protein

(wound response)

JYCHIAN CHEN AND JOSEPH E. VARNER

Department of Biology, Washington University, St. Louis, MO 63130

Contributed by Joseph E. Varner, March 14, 1985

Extensins are hydroxyproline-rich glycopro-ABSTRACT teins associated with most dicotyledonous plant cell walls. To isolate cDNA clones encoding extensin, we started by isolating poly(A)<sup>+</sup> RNA from carrot root tissue, and then translating the RNA in vitro, in the presence of tritiated leucine or proline. A 33-kDa peptide was identified in the translation products as a putative extensin precursor because: (i) it is rich in proline and poor in leucine, and (ii) the message appears to be more abundant when carrot tissue is wounded. From a cDNA library constructed with poly(A)<sup>+</sup> RNA from wounded carrots, one cDNA clone (pDC5) was identified that specifically hybridized to poly(A)<sup>+</sup> RNA encoding this 33-kDa peptide. We isolated three cDNA clones (pDC11, pDC12, and pDC16) from another cDNA library using pDC5 as a probe. DNA sequence data, RNA hybridization analysis, and hybrid released in vitro translation indicate that the cDNA clone pDC11 encodes extensin and that cDNA clones pDC12 and pDC16 encode the 33-kDa peptide, which as yet has an unknown identity and function. The assumption that the 33-kDa peptide was an extensin precursor was invalid. RNA hybridization and DNA sequence analysis indicate that pDC5 is a hybrid clone corresponding to two RNA species. RNA hybridization analysis showed that RNA encoded by both clone types is accumulated upon wounding.

Extensins are hydroxyproline-rich glycoproteins found mainly in the cell walls of dicotyledonous plants (1). These glycoproteins are assumed to play a role in the structure of plant cell walls and may therefore be important in controlling growth, development, and disease resistance. Extensins accumulate in plant cell walls upon wounding (2) and pathogen attack (3) and become insolubilized in cell walls with time (4).

The soluble extensin produced in wounded carrot root tissue has been well characterized both biologically and chemically (5). Chrispeels and Sadava (6) showed that slicing and aerating of carrot tissue enhances the synthesis and secretion of this cell wall component. It has a molecular size of  $\approx$ 86 kDa and consists of one-third protein and two-thirds carbohydrate; the estimated size of the polypeptide is 30 kDa. Six amino acids (hydroxyproline, serine, histidine, tyrosine, lysine, and valine) comprise 95% of the polypeptide backbone (5). The carbohydrate is composed largely of arabinose and a small amount of galactose (7). The carrot extensin is 100% in the polyproline II helical conformation, as determined by circular dichroism spectra; while the deglycosylated molecule is only 50% in the polyproline II helix (7). It appears that the oligosaccharides help in forming and stabilizing the helical structure. Cooper and Varner (4) have found that isodityrosine (two tyrosine residues covalently crosslinked by a diphenyl ether linkage) is formed in carrot cell walls during the insolubilization of extensin. It has been suggested that the soluble extensins are cross-linked by intermolecular isodityrosine residues to form an insoluble network in the plant cell wall (8), although such cross-links have not yet been found.

The determination of the amino acid sequence of extensin is difficult because of the presence of many post-translational modifications and of repetitive amino acid sequences especially imino acids. However, partial amino acid sequences of proteolytic fragments of cell wall-bound tomato extensins have been determined (9). The most significant finding from this work has been the identification of the pentapeptide Ser-Hyp-Hyp-Hyp-Hyp and an intramolecular linkage involving isodityrosine, which has the sequence 1/2 isodityrosine-Lys-1/2 isodityrosine-Lys (10). The similarity of amino acid composition of soluble extensin and the presence of isodityrosine in the insoluble extensin of both carrot and tomato led us to believe that carrot extensins may contain a similar amino acid sequence to that of tomato extensin.

Because the carrot extensins accumulate upon wounding, we decided to construct a cDNA library from wounded carrot root mRNA and isolate the extensin cDNA clones. We herein report the isolation of a cDNA clone that encodes the carboxyl-terminal part of carrot extensin. We also report the isolation of two cDNA clones that encode a proline-rich 33-kDa protein with unknown identity and function.

## **EXPERIMENTAL PROCEDURES**

Materials. Tap roots of carrot (Daucus carota) were obtained from a local merchant and stored at 4°C until use. One-millimeter-thick carrot slices were aseptically prepared with a food processor and aerated in a water vapor-saturated chamber at 25°C for 48 hr. Reverse transcriptase, isolated from avian myeloblastosis virus, was purchased from Life Sciences. Oligo(dT) (12- to 18-mer) and oligo(dT) cellulose were obtained from Collaborative Research. Escherichia coli DNA polymerase I, Klenow fragment, restriction enzymes, T-4 DNA ligase, S1 nuclease, and HindIII linker were obtained from either New England Biolabs, Bethesda Research Laboratories, or Boehringer-Mannheim. Human placental ribonuclease inhibitor was from Promega Biotec (Madison, WI). Ribonucleoside-vanadyl complexes were prepared as described (11). Nucleotides were obtained from P-L Biochemicals. Other chemicals were obtained from Sigma. Isotopes were from New England Nuclear and from Amersham.

**RNA Extraction and Purification.** Two hundred grams of carrot tissue was homogenized in 1 liter of boiling 0.2 M Na borate, pH 9.0/50 mM EDTA/1% NaDodSO<sub>4</sub>/50 mM 2-mercaptoethanol/10 mM ribonucleoside-vanadyl complexes with a Brinkman polytron homogenizer. Solid KCl was added to obtain a final concentration of 20 mM, and the suspension was then held at 0°C for 10 min to precipitate NaDodSO<sub>4</sub>. This solution was subsequently removed by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

centrifugation at  $10,000 \times g$  for 10 min, and solid LiCl was added to the supernatant to 2 M. This solution was stored at 4°C overnight. The precipitated RNA was collected by centrifugation at  $10,000 \times g$  for 10 min. The pellet was resuspended in distilled water and reprecipitated by adding 1/10th vol of 3 M sodium acetate and 2 vol of ethanol. RNA has also been isolated by using the phenol and 100 mM Tris·HCl (pH 9.0) extraction procedure as described (12). The poly(A)<sup>+</sup> RNA was purified by using an oligo(dT) cellulose column as described (13).

In Vitro Translation. Poly(A)<sup>+</sup> RNA was translated by using a wheat germ extract under the conditions described (14) in the presence of [<sup>3</sup>H]leucine or [<sup>3</sup>H]proline. Translation products were analyzed on 10% or 12.5% polyacrylamide gels containing NaDodSO<sub>4</sub> (15), followed by fluorography (16) of the dried gel with Kodak XAR5 film at  $-70^{\circ}$ C. Acetic acid/urea/polyacrylamide gel electrophoresis was performed as described (17).

cDNA Cloning and Analysis. Oligo(dT) (12- to 18-mer) was used to prime the synthesis of complementary DNA from  $poly(A)^+$  RNA that had been size fractionated by sucrose gradient centrifugation (18). After removal of the RNA by alkali treatment. DNA polymerase I was used to synthesize the second strand by the self-priming mechanism (19). Doublestranded cDNAs were trimmed with S1 nuclease, ligated to HindIII linkers, and subsequently digested with HindIII restriction enzyme. These cDNAs were ligated with HindIIIdigested pUC8 (20) and used to transform E. coli JM 83 (21). Transformants were plated on 5-dibromo-4-chloro-3-indolylgalactoside (XGal)/ampicillin plates and white colonies were selected. Another cDNA library was constructed as described by Gubler and Hoffman (22). Briefly, first-strand cDNA is synthesized with  $poly(A)^+$  RNA from wounded carrot and oligo(dT) (12- to 18-mer); second-strand cDNA is synthesized with RNase H and DNA polymerase I; the double-stranded cDNA was tailed with dGTP and cloned into dCTP-tailed puc8 vector. For preliminary characterization, plasmids were isolated from 1.5-ml overnight cultures by rapid alkali denaturation isolation procedure (23). For large scale preparations, plasmids were purified from 1-liter cultures by centrifugation in a CsCl/ethidium bromide gradient. Poly(A)<sup>+</sup> RNA from aerated carrot slices was hybridized with putative cloned cDNAs that were denatured and bound to nitrocellulose disks and eluted for the hybrid-selection translation as described (18). The hybrid-selected RNAs were translated as described above. Colony hybridization was done as described by Maniatis et al. (24).

**RNA Hybridization Analysis.** RNA was separated in 1% agarose gels containing formaldehyde (25) and then transferred to nitrocellulose paper (26). For probes, cDNA inserts were isolated by agarose gel electrophoresis, electroeluted, and nick-translated with <sup>32</sup>P-labeled nucleotides (24). Hybridization of bound RNA was carried out in 40% formamide/3.6× NaCl/Cit (1× NaCl/Cit = 0.15 M NaCl/ 0.015 M Na citrate, pH 7.0)/0.2% NaDodSO<sub>4</sub>/20 mM TrisHCl/5 mM EDTA/3× Denhardt's solution (1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll 400/0.02% polyvinylpyrrolidone) at 42°C for 24 hr and washed with 2× NaCl/Cit/0.2% NaDodSO<sub>4</sub> until the background was acceptable.

**DNA Sequencing.** The pDC5 cDNA insert was recloned into the *Hin*dIII site of the replicative form of mp8 (20), a derivative of the single-stranded phage M13. The singlestranded recombinant phage DNA was used as a template with a universal M13 sequencing primer from New England Biolabs for dideoxy sequencing according to the method of Sanger (27). The inserts of pDC11 and pDC12 were sequenced by the Maxam and Gilbert method (28). The DNA sequence data were stored and analyzed with programs written by W. M. Barnes on a VAX/VMS computer.

## RESULTS

To identify mRNA-encoded extensin, we isolated  $poly(A)^+$ RNA from unwounded and wounded carrot tissue.  $Poly(A)^+$ RNA was translated with a wheat germ extract in the presence of [<sup>3</sup>H]leucine or [<sup>3</sup>H]proline. The translation products separated in NaDodSO<sub>4</sub>/polyacrylamide gels showed a polypeptide band of  $\approx$ 33 kDa rich in proline and poor in leucine (Fig. 1). This polypeptide is present in a much lower quantity in samples from fresh carrot tissue when compared to that from aerated tissue. Extensin as isolated and characterized by Stuart and Varner (5) is (*i*) rich in proline and poor in leucine, (*ii*) a protein component of  $\approx$ 30 kDa, and (*iii*) synthesized actively during wounding. We therefore suspected that this 33-kDa cell-free translation product was the precursor of extensin.

We constructed a cDNA library from size-fractionated RNA that is enriched with the poly(A)<sup>+</sup> RNA encoding this 33-kDa peptide. The cDNA synthesized with oligo(dT) (12- to 18-mer) and size-fractionated RNA was cloned into the *Hind*III site of pUC8 grown in *E. coli* JM83. To identify the cDNA clones encoding this 33-kDa peptide, we did *in vitro* translation of the RNA that was hybrid-selected with cDNA clones. One cDNA clone (pDC5) was isolated that specifically hybridized to the poly(A)<sup>+</sup> RNA encoding the 33-kDa peptide (Fig. 2). The melting temperature ( $t_m$ ) of the cDNA·RNA hybrids was  $\approx$ 90°C under the conditions of these experiments (data not shown).

The cloned cDNA insert from pDC5 was sequenced by the method of Sanger (27). The insert contains 730 nucleotides excluding the 3' poly(A)<sup>+</sup> sequence. In RNA blot hybridization analysis, the pDC5 cDNA insert hybridizes strongly with two poly(A)<sup>+</sup> RNA species (1300 and 1500 nucleotides) present in aerated carrot tissue (Fig. 3), but not in fresh carrot tissue (data not shown). To isolate cDNA clones that represent both RNA species seen in the RNA blot analysis, we constructed another cDNA library by the G-C tailing method. Many cDNA clones with homology to pDC5 were isolated from this cDNA library. Three cDNA clones—pDC11, pDC12, and pDC16—were further characterized. The RNA hybridization analysis (Fig. 3) indicated that pDC11 hybridized to an RNA of 1500 nucleotides.



FIG. 1. Autoradiogram of *in vitro* translation products separated by a NaDodSO<sub>4</sub>/polyacrylamide gel containing 10% acrylamide. One microgram of poly(A)<sup>+</sup> RNA was translated *in vitro* with 1  $\mu$ Ci of [<sup>3</sup>H]proline (lanes A–E and K) or [<sup>3</sup>H]leucine (lanes F–J and L). RNA isolated from unwounded carrot was used in lanes A and F. RNA from three different preparations of wounded carrot was used in lanes B–D, G–I, K, and L. RNA used in lanes E and J was size fractionated to enrich for RNA encoding a 33-kDa protein. The film was underexposed to demonstrate the 33-kDa proline-rich protein band in the gel.



FIG. 2. RNA hybrid-selected by cDNA clones was translated *in vitro* with [<sup>3</sup>H]proline. The translation products were separated by a NaDodSO<sub>4</sub>/polyacrylamide gel containing 12.5% acrylamide. RNA used in lanes A, B, C, and D was selected by pDC5, pDC12, pDC16, and pDC11, respectively. Poly(A)<sup>+</sup> RNA isolated from wounded carrot was used for the hybrid selection experiment. Two micrograms of poly(A)<sup>+</sup> RNA isolated from wounded carrot was used in lane E. Lane F has labeled protein molecular size markers.

The insert DNA of pDC11, pDC12, and pDC16 was sequenced by the Maxam and Gilbert method (28). The DNA



FIG. 3. RNA hybridization analysis of  $poly(A)^+$  RNA isolated from wounded carrot hybridized to nick-translated probes. Five micrograms of RNA was separated by a 1% agarose gel containing formaldehyde, transferred to a nitrocellulose paper, and hybridized to nick-translated insert DNA of pDC11 (lane A), pDC5 (lane B), and pDC12 (lane C). sequence of pDC11 has 386 nucleotides identical to the 3' part of pDC5 sequence (345–730), except for 13 nucleotides at the 5' ends. This sequence encodes a single open reading frame containing 43 amino acids (Fig. 4). The predicted amino acid sequence includes four Ser-Pro-Pro-Pro-Pro repeats, one Tyr-Lys-Tyr-Lys, and one Thr-Pro-Val. These sequences indicate that pDC11 is an extensin cDNA clone (see *Discussion*).

We made an attempt to identify the translation product that corresponds to the extensin clone pDC11. While a NaDod-SO<sub>4</sub>/polyacryamide gel system failed to resolve any putative extensin protein, an acetic acid/urea gel system did resolve several protein bands (Fig. 5). The presence of several bands in acetic acid/urea gel system may reflect minor heterogeneity of extensins or cross-hybridization with other mRNAs in the hybrid selection.

Upon compiling the sequences from pDC5, pDC12, and pDC16, we found that the sequences of pDC5 and pDC12 are identical from nucleotide 281 to 340. Also, it was found that, except for a few single base changes (Fig. 6), pDC16 overlaps the sequence formed by the combination of pDC5 and pDC12 from nucleotide 130 to 622. Therefore, we conclude that pDC5, pDC12, and pDC16 represent a single continuous

	10	20	30	40	50	60
CCATTT	TAGGAGTTGC	GATTTTCGCA	GCTCCATCC	TAGCCGATT	TCATTCTCA	CCCG
laiter	eucryvarai	allerneala	alarroseri	Leualaaspri	len1ssern1s	spro
	70	80	90	100	110	120
CAATTC	ACAAGCCTCC	CGTTTATACA	ACCACCCGTT(			CCT
rollen	ISLYSFFOR	ovallyrin	rrorrovali	iislyspropi	COLLEHISLYS	srro
	130	140	150	160	170	180
CCCGTTT	ATACACCACC	CGTTCACAA	CCACCAGTT		CCGTTCACAA	ACCA
rovall	yrinreroei	ovarmistys	sproproval	lyr inreroei	rovathisLys	srro
	190	200	210	220	230	240
CTTCCC	AATATAAACO	TCCGGTGGA	AGCAACCAAC	TCCGTGACTG	AAGATCATTA	СССТ
roserG	luTyrLysPr	roprovalgi	IAlaThrAsh:	SerValThrG.	luAspHisTy	rPro
	250	260	270	280	290	300
ATTCACA	AGCCACCAG	IGTACAAACC/	ACCTGTTCAG	AAGCCAGCAC	CAGAGCACAA	AČCA
lleHisL	ysProProVa	alTyrLysPro	ProValGln	LysProAlaPi	roGluHisLy	sPro
	310	320	330	340	350	360
CCCGTTC	ACAAGCCACO	CAATTCACAA	ACCACCGGTT	CACACACTAG	ITTACAAGTA	CAAG
ProValH	lisLysProPı	rolleHisLys	sProProVall	HisThrLeuVa	alTyrLysTyı	rLys
		pi	Proi	ProThrPro.	•	
	370	380	390	400	410	420
SerProF	roProProM	etHisSerPro	ProProProPro	JIIIACICICO ValTvrSerPi	CACCACCACCO	CAAA
				un yr oer r		51.95
	430	440	450	460	470	480
CATCACI Highigi	ACTCCTATA(	CGTCACCTCC	CCTCCTCAC	CACTACTAAT!	AAAAACTCTC	CCTA
	.yr sei ryr n			lisiyi		
	490	500	510	520	530	540
AAGGACA	CTGATGGAG	GCCAACTTAA	GAAGATGAAG	TAAAATAATG	GCTTGCACGA	TGAT
	550	560	570	580	590	600
TTTGGCC	<b>STTTTTATAA</b>	ATATGTTGAT	CGAATATGAT	ATTTTTGTTG	TATTCTAGTA	TGGG
	610	620	620	610	650	660
TCATTTO	CTTTTTGTT	TGCGAATGAA	TAACATTTAC	ATATGCATGT.	AGCATCGGGG	AACC
ተርጉተር እ	670 °CTCTAGAAA	680 TAGAGGTTTC	690 TTATTCTCCC	700	710 ******	720
ICTION	JOI UI AUAAAA		THIUUUUU	AT THANGORY	IIAIAIGAAI.	AAGI
	730	740	750	760	770	780
ATTTCT	STTA					

FIG. 4. DNA sequence of pDC5 and pDC11. The insert DNA of pDC5 was subcloned into M13 phage vector mp8 and sequenced according to the Sanger method. The insert DNA of pDC11 was sequenced by the Maxam–Gilbert method. pDC11 is identical to pDC5 from position 345 to 730 but has 13 nucleotides preceding this homology that are unique. The predicted amino acid sequence is also shown.



FIG. 5. RNA hybrid-selected by cDNA clones was translated *in vitro* with [<sup>3</sup>H]proline. The translation products were separated by acetic acid/urea/polyacrylamide gel containing 12.5% acrylamide. RNA used in lanes B, C, and D was selected by pDC11, pDC12, and pDC5, respectively. Two micrograms of poly(A)<sup>+</sup> RNA isolated from a wounded carrot was used in lane E. Lane A is a control without exogenous RNA. Arrows indicate the most prominent bands.

sequence of 832 nucleotides. In this DNA sequence, there are several repetitive nucleotide sequences with only minor differences. The longest repeat unit is 54 nucleotides long.

This 832-nucleotide sequence has a single open reading frame encoding 211 amino acids (Fig. 6). The predicted peptide sequence contains 24 repetitive units of X-Y-Pro-Pro. The single nucleotide changes between pDC16, pDC5, and pDC12 are always in the wobble position. The amino acid codons CCA and CCT for proline were preferentially used.

To identify the peptide encoded by this cDNA sequence, we did *in vitro* translation with RNA selected by hybridization to pDC12 or pDC5. The translation products are the same 33-kDa proline-rich and leucine-poor peptide when separated by NaDodSO<sub>4</sub>/polyacrylamide gels (Fig. 2). This indicates that pDC12, pDC16, and the 5' part of pDC5 are encoding this 33-kDa protein. Because the size of the RNA hybridized to pDC12 is  $\approx$ 1300 nucleotides, and in the predicted peptide sequence does not contain methionine, we suggest that the peptide sequence represents only the carboxyl-terminal part of the protein. RNA hybridization, DNA sequence analysis, and the *in vitro* translation of hybridselected RNA indicate that the insert of pDC5 is a hybrid cDNA resulting from a cloning artifact, most likely due to a ligation step.

It has been demonstrated that extensins accumulate in plant cell walls upon wounding (2). However, it is not known what effect wounding has on the extensin RNA. We examined this by hybridization of RNA from unwounded or wounded carrot root with the cDNA probes (Fig. 7). The extensin RNA was detected by probing with pDC11, and the 33-kDa protein RNA was detected with pDC12. The results indicated that both kinds of RNA accumulate upon wounding.

## DISCUSSION

The cell wall-bound hydroxyproline was first investigated in the 1960s. However, the insolubility of the peptidyl hydroxyproline, called extensin by Lamport, coupled with

Proc. Natl. Acad. Sci. USA 82 (1985)

pDC5	I	340	)					
pDCI2		2 <u>82</u>			832			
pDC16	85			622				
	•							
CODING				— (211 ar	nino acids)			
GCCATTTT AlaIleLe	10 AGGAGTTG uGlyValA	20 CGATTTTCGC lallePheAl	30 AGCTCCATCC aAlaProSer	40 CTAGCCGATT LeuAlaAspP	50 TTCATTCTCA heHisSerHi	60 CCCG sPro		
CCAATTCA	70 CAAGCCTCO	80 CCGTTTATAC	90 ACCACCCGTT	100 Cacaaaccac	110 Caattcacaa	120 GCCT		
ProIleHi	sLysProPi	roValTyrTh	rProProVal	HisLysProP	rolleHisLy	sPro		
1	30	140	150	160	170	180		
ProValTy	rThrProPi	roValHisLy	sProProVal	TyrThrProP	roValHisLy	acca sPro		
1 CCTTCCGA	90 Atataaaco	200 CTCCGGTGGA	210 AGCAACCAAC	220 ICCGTGACTG	230 Aagatcatta	240 СССТ		
ProSerGl	uTyrLysPi	roProValGlu	G uAlaThrAsn	C SerValThrG	luAspHisTy	rPro		
29 ATTCACAA IleHisLy	50 GCCACCAGI sProProVa	260 IGTACAAACCA alTyrLysPro	270 ACCTGTTCAG ProValGlni	280 AAGCCAGCAC LysProAlaPi	290 CAGAGCACAAA roGluHisLy:	300 ACCA sPro		
3 CCCGTTCA ProValHi	10 CAAGCCACO sLysProPi	320 CAATTCACAAA colleHisLys	330 ACCACCGGTT( sProProVall	340 CACAACACTCO HisAsnThrPi	350 CCTCCGTAAC roSerValThi	360 TGAT rAsp		
3 GATCATTA AspHisTy	70 CCCGGCCC/ rProAlaHi	380 ATCCTATTCA IsProlleHis	390 FAAACCACAA SLysProGlni	400 CCGATTCACA( ProIleHisA)	410 GGCCACCGGT rgProProVal	420 ICAC LHIS		
4 AAGCCACC LysProPro	30 AACAGAGC <i>i</i> oThrGluHi	440 ATAAACCTCCC IsLysProPro	450 GGTTCACGAG ValHisGlui	460 CCAGCGACAG ProAlaThrG	470 AGCACAAGCCO LuHisLysPro	480 GAGC Ser		
4 CCAGTGTA	90 CCAACCACC	500 CGAAGACTGAO	510 GAAGCCTGTGG	520 CCAGAACATA	530 Agecaccacad	540 CTTG		
ProValTy	rGlnProPr	roLysThrGlu	LysProVall	ProGluHisLy	/sProProHis	sLeu		
5 CCACCAAT	50 CGTCGTGAC	560 GCCACCGCC1	570 FACACATAAGO	580 ССТААТССТСО	590 Catacggtcad	600 CCAC		
ProProIl	eValValAr	rgProProPro	oThrHisLysl	ProAsnProPi	roTyrGlyHis	sHis		
6 CCTAGACA ProArgHi	10 CCCTCCAGI sProProVa	620 FTGAGAATACO AlGluAsnThr	630 CGGAAACTAAO CGlyAsn	640 Gacgtcatgt/	650 Agtgtaatta(	660 Caat		
6' Aaataaga	70 ACTGGTGA <i>i</i>	680 MAATAATAATO	690 STTACGAGTG	700 FATGTATCTAT	710 FAAGCTCATT	720 ACCT		
7: CTTATGTT	30 ICGGCCAA	740 AATAGGTTATO	750 Gagcaagtag:	760 FACTATATGCI	770 Igtaataaato	780 CAGT		
71 GTATGAGT	90 FATCCATGI	800 FATTACGCAT	810 Itactgttati	820 Atagcaaaagi	830 FTATAC	840		
FIG. 6	A diagr	am of overl	anning sea	ience of nD	C5. nDC12	and		

FIG. 6. A diagram of overlapping sequence of pDC5, pDC12, and pDC16 (*Upper*), and the compiled DNA sequence from pDC5 and pDC12 along with the predicted amino acid sequence (*Lower*). The insert DNAs of pDC12 and pDC16 were sequenced by the Maxam-Gilbert method. Nine nucleotides of pDC16 are different from the compiled DNA sequence as indicated. Only one of these nucleotides (at position 94 of the compiled sequence) will change the encoded amino acid sequence (from valine to isoleucine).

the unusual amino acid composition have made their isolation and characterization a difficult task. Lamport (9) isolated several small repetitive peptides after acid and trypsin hydrolysis of tomato cell walls. All of the peptides characterized contain the sequence Ser-Hyp-Hyp-Hyp-Hyp. Also, isodityrosine residues present in the insoluble tomato extensin were found in two of the fragments (10). One



FIG. 7. RNA hybridization analysis of RNA isolated from wounded (lanes B and D) and unwounded (lanes A and C) carrot hybridized to nick-translated probes. Twenty-five micrograms of total RNA was separated by a 1% agarose gel, transferred to nitrocellulose paper, and hybridized to nick-translated insert DNA of pDC11 (lanes A and B) and pDC12 (lanes C and D). The film was overexposed to demonstrate that only a trace of extensin mRNA is present in unwounded carrot.

fragment has the sequence Ser-Hyp-Hyp-Hyp-Hyp-Val-1/2isodityrosine-Lys-1/2 isodityrosine-Lys and the two tyrosine residues in this fragment are linked to form an intramolecular isodityrosine (10). Isodityrosine is also found in carrot (4) and many other plant cell walls (8) and is believed to be characteristic of extensins. Recently, Smith *et al.* (29) were able to isolate soluble extensins from tomato tissue culture by eluting the cells with salts. They identified two different extensins in these eluates. The amino acid composition of tomato soluble extensin P1 is similar to the soluble extensin isolated from carrot (7). From these data, we believe that carrot and tomato extensins contain similar peptide sequence.

From cDNA libraries constructed with poly(A)<sup>+</sup> RNA of wounded carrot, we have isolated several cDNA clones. The cDNA clone pDC5 was identified by its specific hybrid selection of an RNA encoding a 33-kDa protein that is rich in proline and poor in leucine. However, our data indicate that the insert of pDC5 is a hybrid cDNA resulting from a cloning artifact. The 5' part of pDC5 insert overlaps with the cDNA clones pDC12 and pDC16 and encodes a 33-kDa protein. Although only a partial peptide sequence can be derived from the cDNA sequences for the 33-kDa protein, it is evident that the amino acid sequence does not have the characteristic amino acid sequence Ser-Pro-Pro-Pro-Pro, which was expected by analogy with the tomato extensin. Therefore, clones pDC12 and pDC16, although they encode a prolinerich protein of 33 kDa, probably do not encode extensin. It will be interesting to further characterize this protein and determine if any of the prolines are hydroxylated.

The 3' part of the hybrid clone pDC5 that overlaps with pDC11 encodes a peptide containing Ser-Pro-Pro-Pro-Prorepeats and Tyr-Lys-Tyr-Lys. Moreover, a carrot genomic clone pDC5A1, which has a homologous sequence to pDC5 and pDC11, encodes the same peptide containing 25 Ser-Pro-Pro-Pro-Pro sequences (unpublished observations). It also has a sequence of Ser-Pro-Pro-Pro-Pro-Thr-Pro-Val-Tyr-Lys, which was found in tomato extensin (9). Based on the evidence presented here, we believe that a part of pDC5 and of pDC11 encodes the precursor for the extensin glycoprotein.

Extensin is accumulated in plant cell walls upon wounding

(2), and the RNA hybridization data (Fig. 3) show that extensin mRNA also accumulates upon wounding. We suggest that the increase in extensin synthesis is a part of the wound response and is regulated at the level of transcription. We do not know whether the 33-kDa proline-rich protein is accumulated during wounding, but the RNA for this protein is accumulated in response to wounding. The 33-kDa protein may be involved in the response of wounding stress in carrots. Because both extensin and the 33-kDa protein RNA are accumulated during wounding, the mechanism for the gene regulation of these two proteins should be interesting.

The DNA sequence of the 33-kDa protein cDNA showed that there are several repetitive nucleotide sequences with some single base changes at the wobble positions. We suggest that intragenic duplications may have occurred during the evolution of this gene.

We thank Steve Janovec for his excellent editorial assistance. This work was supported by National Science Foundation Grant PCM8104516, Department of Energy Grant F602-84ER13255, and an unrestricted grant from the Monsanto Company to J.E.V.

- Lamport, D. T. A. & Catt, J. W. (1981) in *Plant Carbohydrates II:* Encyclopedia of Plant Physiology, New Series 13B, eds. Tanner, W. A. & Loewus, F. A. (Springer, Berlin), pp. 133-165.
- Chrispeels, M. J., Sadava, D. & Cho, Y. P. (1974) J. Exp. Bot. 25, 1157-1166.
- Esquerre-Tugaye, M. T. & Lamport, D. T. A. (1979) Plant Physiol. 64, 314-319.
- 4. Cooper, J. B. & Varner, J. E. (1983) Biochem. Biophys. Res. Commun. 112, 161-167.
- 5. Stuart, D. A. & Varner, J. E. (1980) Plant Physiol. 66, 787-792.
- Chrispeels, M. J. & Sadava, D. (1974) in Macromolecules Regulating Growth and Development, eds. Hay, E. D., King, T. J. & Papaconstantinou, J. (Academic, New York), pp. 131-152.
- 7. Van Holst, G. J. & Varner, J. E. (1984) Plant Physiol. 74, 247-251.
- 8. Fry, S. C. (1982) Biochem. J. 204, 449-455.
- Lamport, D. T. A. (1977) in Recent Advances in Phytochemistry, eds. Loewus, F. A. & Runeckles, B. C. (Plenum, New York), Vol. 2, pp. 79-115.
- 10. Epstein, L. & Lamport, D. T. A. (1984) Phytochemistry 23, 1241-1246.
- 11. Berger, S. L. & Birkenmeier, C. S. (1979) Biochemistry 18, 5143-5149.
- 12. Haffner, M. H., Chin, M. B. & Lane, B. G. (1978) Can. J. Biochem. 56, 729-733.
- Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
- 14. Roberts, B. E. & Paterson, B. M. (1973) Proc. Natl. Acad. Sci. USA 70, 2330-2334.
- 15. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
- 16. Jen, G. & Thach, R. E. (1982) J. Virol. 43, 250-261.
- 17. Spiker, S. (1980) Anal. Biochem. 108, 263-265.
- Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. & Sidman, J. G. (1981) Proc. Natl. Acad. Sci. USA 78, 2253-2257.
- Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) J. Biol. Chem. 253, 2483-2495.
- 20. Vieira, J. & Messing, J. (1982) Gene 19, 259-268.
- 21. Enea, V., Vovis, G. F. & Zinder, N. D. (1975) J. Mol. Biol. 96, 495-509
- 22. Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269.
- 23. Ish-Horowicz, D. & Burke, J. F. (1981) Nucleic Acids Res. 9, 2989-2998.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) Biochemistry 16, 4743-4751.
- Thomas, P. S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
  Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 28. Maxam, A. M. & Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
- 29. Smith, J. J., Muldoon, E. P. & Lamport, D. T. A. (1984) Phytochemistry 23, 1233-1239