# Sequence Analysis and Comparison of Avocado Fruit and Bean Abscission Cellulases<sup>1,2,3</sup>

# Mark L. Tucker\* and Stephen B. Milligan<sup>4</sup>

U.S. Department of Agriculture, Agricultural Research Service, Plant Molecular Biology Laboratory, Room 207, Building 006, BARC-West, Beltsville, Maryland 20705

#### ABSTRACT

A 1700 nucleotide cDNA clone for a bean (Phaseolus vulgaris cv Red Kidney) abscission cellulase (endo-(1,4)- $\beta$ -D-glucanase) has been identified and sequenced. This cDNA clone contains a 1485 nucleotide open reading frame which includes coding sequences for a putative signal peptide and mature protein. The nucleotide and deduced amino acid sequences for the bean abscission cellulase are compared to the previously reported sequences of an avocado fruit ripening cellulase. Optimal alignment of these sequences shows 64% and 50% identically matched nucleotides and amino acids, respectively. Analysis of the deduced amino acid sequences for the mature bean and avocado cellulases indicates that these two proteins share similar molecular weights, position of cysteine residues, and hydropathic character, but have very different isoelectric points and glycosylation. Genomic blot data suggest that the avocado fruit cellulase belongs to a small gene family, whereas the bean abscission cellulase appears to be encoded by a single gene or a few very closely related genes.

Cellulase activity, as defined by the ability to degrade carboxymethylcellulose, has been measured in a variety of plant organs, tissues, and developmental stages, *e.g.* abscission (11, 20), fruit softening (1, 2), cell elongation (4), lateral root initiation (16), and anther development (21). Avocado (*Persea americana* cv Hass) fruit is quite extraordinary with regard to the large quantity of cellulase that is synthesized during fruit ripening (1, 15). *In vitro* translation of avocado mRNA suggests that the message for immunoprecipitable cellulase accumulates to as much as 2% of the translatable mRNA in the mesocarp of ripe soft fruit (23). It was from this source that the first plant cellulase cDNA clone was identified (5).

An avocado fruit cellulase cDNA, pAV363 (24), was used to identify a cDNA clone, pBAC1, for a considerably less abundant bean (*Phaseolus vulgaris* cv Red Kidney) leaf abscission cellulase (25). The previously characterized 595 nucleotide abscission cellulase clone, pBAC1, is a partial clone of an approximately 2.0 kb<sup>5</sup> cellulase mRNA (25). We have subsequently identified and sequenced a 1700 bp abscission cellulase cDNA clone from bean which encodes a putative signal peptide and the entire mature protein. The deduced amino acid sequence for the bean abscission cellulase is compared to that of the avocado fruit cellulase to define sequence and structural similarities and differences between these two plant cellulases. Moreover, the sequence comparison will aid in the identification of related cDNA and genomic clones which, in turn, will contribute to an understanding of the role of cellulases in plant development. Genomic Southern blot results for both cloned probes, pAV363 and pBAC10, are analyzed with regard to the potential relationship of different genes within a cellulase gene family (24).

#### MATERIALS AND METHODS

### **Cloning of cDNA**

Abscission zone tissue was cut from bean (Phaseolus vulgaris) explants as previously described (25). Plants were grown for 12 d, stem and petiole explants with primary leaves removed were exposed to 25  $\mu$ L/L ethylene in air for 48 h, and a 2 mm cross section of the lower abscission zone was harvested and frozen in liquid nitrogen. Abscission zone tissue was mixed with dry ice, pulverized in a coffee grinder, and stored at  $-80^{\circ}$ C. Pulverized abscission zones (10 g) were suspended in 50 mL of polysome extraction buffer (0.5 M sucrose, 200 mм Tris [pH 9.0], 400 mм KCl, 35 mм MgCl<sub>2</sub>, 25 mM EGTA, 5 mM 2-mercaptoethanol), and polysomal RNA was prepared as described by Jackson and Larkins (12). Poly(A)<sup>+</sup> RNA was purified by oligo(dT) chromatography (23) and cDNA synthesized using the procedure of Gubler and Hoffman (9). The cDNA library was cloned by insertion into the *Eco*RI site of  $\lambda$ GT10 as described in the protocol for Amersham's cDNA cloning system. The phage library was screened with the partial length pBAC1 bean abscission cellulase cDNA (25) and a 5' PstI cDNA fragment of pAV363, the avocado fruit cellulase cDNA (24).

## **Nucleotide Sequence**

Single- and double-stranded DNAs were sequenced by the dideoxy chain termination method of Sanger *et al.* (19) as described in the Sequenase protocol (United States Biochem-

<sup>&</sup>lt;sup>1</sup> Supported in part by National Science Foundation grant DCB-8310460 and U.S. Department of Agriculture grant 8900650.

<sup>&</sup>lt;sup>2</sup> In memory of Dr. Jacob B. Biale.

<sup>&</sup>lt;sup>3</sup> The nucleotide sequence of the bean abscission cellulase cDNA has been deposited in the GenBank sequence library, accession number M57400.

<sup>&</sup>lt;sup>4</sup> Present address: Department of Biochemistry and Biophysics, University of California, Davis, CA 95616.

<sup>&</sup>lt;sup>5</sup> Abbreviations: kb, kilobase; kbp, kilobase pair; bp, base pair; PCR, polymerase chain reaction; pl, isoelectric point.

ical Corp.). The sequence was fully confirmed by obtaining overlapping sequence data from both strands of several cDNA clones (pBAC1, pBAC3, pBAC5, pBAC10, and PCR fragment clone). Both endonuclease restriction fragment subclones and synthesized oligonucleotide primers were used for sequencing the 1700 nucleotides of cDNA.

#### **Genomic DNA Hybridization**

DNA was isolated as previously described (24) from frozen powdered mesocarp of avocado fruit and young unexpanded leaves of bean. Ten micrograms of endonuclease digested DNA were loaded into each lane of a 0.8% agarose gel, electrophoresed, then transferred and fixed with UV irradiation onto Hybond-N (Amersham nylon membrane) (17). The cDNA inserts used as probes were separated from the vectors and purified as described by Langridge *et al.* (14). Labeled probe was removed from blots prior to rehybridization using 0.4 M NaOH at 45°C as described in the Amersham protocol for Hybond-N membrane. Hybridization and washing conditions are described later.

#### **Sequence Analysis**

Optimal alignment of DNA and amino acid sequences and analysis of deduced polypeptides were made using the Genetics Computer Group (GCG) sequence analysis software package (7).

## RESULTS

## Identification of cDNA Clone

A bean abscission-specific cDNA library was cloned by insertion into the *Eco*RI site in  $\lambda$ GT10. The cloned library was screened with the previously reported bean abscission cDNA clone pBAC1 (25). The set of clones isolated through this screening were all of approximately equal length, 800 bp. However, a full-length cDNA was expected to be approximately 2.0 kbp based on Northern blot analysis (25). The same bean cDNA library was then secondarily screened with a 5' restriction fragment of the avocado fruit cellulase cDNA (5' PstI fragment of pAV363). This second screening yielded clones of mixed length from 200 to approximately 900 bp that did not cross-hybridize with the first set of clones. The longest clones from the first and second screenings of the bean cDNA library were sequenced and found to have sequence similarity with the 3' and 5' ends, respectively, of the avocado cellulase cDNA. As part of the cloning procedure, EcoRI methylase was used to protect internal cDNA restriction sites from being digested when EcoRI endonuclease was used to remove excess EcoRI linkers from the ends of the cDNA. It occurred to us that there might be one or more EcoRI site(s) in the middle of the bean abscission cDNA that were not protected by methylation with EcoRI methylase resulting in our cloning separate 3' and 5' fragments of the full-length cDNA. To determine if there were indeed one or more internal EcoRI sites, we prepared oligonucleotides complementary to unique sequences in the respective 3' and 5' bean clones and used the PCR method (18) on abscissionspecific cDNA to amplify the intervening sequence. The PCR

fragment was then blunt end-ligated into the *Hin*cII site of the plasmid vector  $T7/T3\alpha 18$ , cloned in DH5 $\alpha$ , and the cloned fragment sequenced. The sequence of the PCR fragment showed two internal *Eco*RI sites separated by 48 nucleotides. The DNA sequence on either side of this internal 48 nucleotide segment was identical to the 3' and 5' bean cellulase clones indicating that the two partial clones and the PCR product were encoded by the same gene.

We then generated a clone, pBAC10, that encompasses the sequences of all the partial clones. This was done by: (a) using PCR to extend the 3' and 5' overlapping ends of gel purified inserts from pBAC5 and the cloned PCR fragment, respectively, (b) PCR amplification of the extended fragment using oligonucleotides complementary to each end of the extended product, and (c) inserting the extended amplified product into the *Hin*CII site of the vector T7/T3 $\alpha$ 18. The endonuclease restriction map and nucleotide sequence derived from sequencing the pBAC1, pBAC3, pBAC5, pBAC10, and PCR fragment clones is shown in Figure 1.

## Alignment and Analysis of Nucleotide Sequences

Optimal alignment of the bean and avocado cellulase nucleotide sequences shows 64% sequence identity between their respective open reading frames (data not shown). The 3' end of the bean nucleotide sequence (Fig. 1) is A/T rich and contains a few sequences that resemble the polyadenylation signal consensus sequences (AATAAA); however, no consensus sequence is localized within the most common position of 15 to 23 nucleotides from the end of the mature message.

## Alignment and Analysis of Amino Acid Sequences

The largest open reading frame beginning with an ATG (methionine) in the bean cDNA sequence results in a polypeptide of 495 amino acids (Fig. 1). The molecular mass of this predicted polypeptide is 54.4 kD. This predicted value is comparable to the 51 kD value determined by SDS-PAGE of the *in vitro* translation product of hybrid-selected mRNA (25).

Both avocado and bean cellulases are secreted proteins. Both deduced polypeptides have a predicted signal peptide characteristic of secreted proteins (26). The amino acid sequence of the signal peptides is not conserved; however, there are several regions within the aligned sequences of the mature polypeptides that are conserved (Fig. 1). Optimal alignment of the deduced polypeptide sequences (Fig. 1) shows 50% identically matched amino acids. Optimal alignment furthermore shows there to be several small deletions and insertions. The longest insertion in the bean sequence (or deletion from the avocado sequence) is at the amino terminus of the mature peptide. The functional or structural significance of this additional amino terminal sequence in the bean cellulase is unknown.

The mol wts of the mature polypeptides for the avocado and bean cellulases differ by less than 1% (Table I), and the positions of all six cysteine residues in the mature bean sequence align with cysteine residues in the avocado sequence (Fig. 1). Furthermore, the two cellulase polypeptides have very similar hydropathy plots (Fig. 2). These common char
 Kpni
 Ssti
 EcoRi
 EcoRi
 Bglii

 0
 bp
 pBAC10
 1700 br

 pBAC5
 pBAC3

 100 bp
 PCR fragment

 100 bp
 pBAC1

 CTCATACTCATATATCATGGGGTACCATTCGGTGTTCATTGCTGTGTTCTTATGGAGTTCAATGGTTTGTCACAATGGATTAGCCATG
 8

 CTCATACTCATATATCATGGGGGTACCATTCGGTGTTCATTGCTGTGTCAATGGATTAGCCATG
 8

 DV0CDD0
 V
 0
 0
 0
 1
 2

C	TCA	TAC	TCAT BEJ	iata N	TCA:	rGGG G	GTI Y	ACCI H	ATTC Ş	GGT V	GTT F	CAT I	TGC Å	TGT( V	GTT F	CTI L	'ATG W	GAG. S	rtci S	NATO M	GGTI V	rtg: C	ICA H	CAA' N	IGG	ATT:	AGC	CATG	88
		AV	OCAL	ю	M	D	С	S	5	₽	L	s	L	F	H	Ł												L	13
ATG M	GAT D	GAT D	GGG7 G 1		TGA	CGAG	CTC	CTC	AGG	ငင္ရင	ACC	TAA	TTA	TGA	CŢA	тçc	AGA	TĢC	IÇT]	rçco	C NN	včci	[ĂT]	ŢŢ	GTT	TTT	IGA	AGGA	178
L	v	с	TV	 / М	 1 V	ĸ	c	c	s	- A	- *s	D	т.	н Н	ţ	g	ĭ	Ĵ	Ĩ	<b>*</b>	Ĵ	A.	ţ	Ţ	Ĩ	Ĩ	ļ	ļ	54
CAA	čec	TCA	GGG	AGT	TGC	CTC	AAC	ic cz	AAG	AGT	GAA	- G <u>T</u> G	GAG	GGAI	AGA	СТС	CGC	ICT(		GA:	r GGC	3 2881	L LCT		г Алл	r CGT(	e Gaa	ТТТС	43 268
Ĭ	I I			. ц . Ц	· P	S _	S	Î	R	v	ĸ	W	R	E	D	S	λ	L	8 	D	Ğ	ĸ	Ľ	Q	N	Ÿ	N	Ļ	84
Q ATG	K GGA	5 ( 661)	G .К Гаст	. Г Э.СС	. Р Зтс(	T TCC	N TCI	ע ג גרידי א	R	L	T	W	R	G	D	\$ 777	G	Ĺ	Ś	Ď	Ġ	S	S	Y	H	Ϋ́	D	Ĺ	73
M	Ģ	G	ŶŶ	, D	À	Ğ	D	Ň	V	K	F	Ģ	W	P	M	GGC À	F	S	T T	S	L L	L L	S	rtg( W	JGCI	IGC: A	IGT V	AGAG Ē	358 114
v	Ġ	Ġ	Ϋ́Υ	с р	Å	Ġ	Ъ	Ņ	L	ĸ	F	ę	L	P	Ņ	¥	F	T	Ŧ	т	M	Ł	λ	Ŵ	G	I	I	Ē	103
TAC Y	GAG. E		GAGA E I	TAT	CTTC	CGI V	'GAA N		GCT L	TGG: G	TTA: Y	ICT L	CCA	AAG1	IGC A	TAT I	CCC	TTGO	GCI	igcz	AGAC	TT	ATC	STT(	SAG	AĞCI	AÇA	таст	448
F	G	c :	LM	P			E	Į	v	E	N	A	R	λ	Ā	L	Î	Ï	s	т	Ĩ	- Y	т.	Ĩ	ĸ	Ĵ	9	Ţ	131
TCC	CCA	ACC	À	ccç	TCŢA	TĂC	ACA	GĢT	тçG	AGA:	rga	AĂY	ଟ୍ଟେଟ	IGAI	C.M	CAA	TTG	TIGO	GAG	ccc	- 		GAC	ATC	GGA:	TACI	ICC.	AAGG	535
2	т	- c 1	. 1 . e	Į	Ţ	1 17	Ĭ	ľ	Ĭ	7	5	ļ	Ĩ	Ĭ	1	N	ľ	<u>N</u>	<u>r</u>	R	P 	E 	D	M	P	Ŧ	P 	R 	173
ACA	GTC	TAT	AAGA	TAG	ATGO	Таа	Y TTC	TCC	AGG	27C	P TGBI	N AGTI	A CCC2	ע הכטו	H Ga	.К Ста		W	e E	R	P	E	D	M	D	T	P	R	161
T	Ŷ	Y I		D	À	N	S	P	Ğ	T	E	Ÿ	A I	A I	E	Ŷ	A A	Å	A	Ļ	S	À	A	S	I	V	F	K	203
N	Ŷ.	Ý 1	K V	S	T	Q	N	Þ	Ġ	S	D	Ý	Å	Á	Ė	T	Å	Å	Å	ŗ	A	Å	Å	ś	ł	v	F	G	191
AAA K	ATA I l	GATO D 1	SCCA N K	AGT/	ATTC Ş	CTC S	CAC T	GTT Ļ	ATT/ Ļ	AAGO S	CCA1 H	rtci S	AAAJ K	ATCA S	L	STT F	TGA: D	FTTT	'GCA A	GAC D	AAG K	AAC N	AGA R	.GGI G	ITCI S	TAC Y	TC(	CGGT	715 233
D	S I		3 S	Ŷ	s	т	ĸ	ŗ	Ł	н	т	A	v	ĸ	v	F	E	F	Å	l D	Q	Y	 R	ļ	l s	Î Y	Ì	D	221
TCT	• • •	• • •	• • • •	• • •	<u>T</u> G	င္ရင္ရင	ATT	CŢA	CIG	CŢCO	5ŢAC	TC	AGGI	TAC	CAC	GA	TGAI	<b>\</b> ÇTG	ÇTA	TGG	GCT	GCT	GCT	TGG	CTT	TAC	-	GCG	790
Ĩ.	τ. (		 . v	v	Ĭ	ļ	Ĩ	Ĵ	Ĭ	Į	ţ	Ĩ	Ĭ	ţ	¥ N	ĭ	Ĩ	Ĭ	ţ	<u>.</u>	A 2	1	A -	Ĩ	Ī	¥ 	ĸ	Â	258
AGT	GGA	GAN	GCA	AGT	ATCT	GAG	CTA	CAT	CAT	AAGI	TAAC	S CA	G	1	N	ເດດ	e Ntga		CJG	W 3/73	G GTC	А Т.С.Т.	S Crr	W ጥጥር	Р	H TCC	R	A	251
s (	G	2 5	S K	Y	L	S	Ŷ	I	I	S	N	Q	•	•		Ğ	Ŵ	S	Q	T	v	Š	E	F	S	Ŵ	D	N	285
S (	ຊີ		S	_ Ÿ	M	T	Ý	İ	Q	Ś	Ņ	G	H	T	L	Ġ	A	D	D	D	D	Y	S	f	s	Ŵ	b	D	281
K 1	F V		ι α α	Q	AGAC T	ATT. Ļ	ACT.	AAC T	E	GGAJ E	F	Y Y	GGI	GCC	AAG K	;	. AAJ K	LGAC D	TTG Ļ	GCC A	AAG K	ATT I	aag K	АСІ Т	GAC D	GCT	GAC E	STCA S	958 314
ĸ 1	R 1	7 6	; Т	K	V	Ł	ŗ	S	ĸ	G	F	L	Q	D	R	I	E	E	Ł	Q	L	Y	ĸ	v	H	т	D	N	311
TTT	ATA: I (		CTG	TGAT M	IGCC P	AGG	AAG	TAA N	CTC	CGJ R	CAC	ATT I	raar K	ACA	ACI	10C	IĞCI	Gec	CTT	CTG	TTT	ACC	AGA	GAT	AGT	AGC	AAT	ŢTG	1048
Y			L	I	ļ	Ĭ	т	s	Ĩ	F	Ţ	2	0	v	Î.	Į.	Į	Į	Ĩ	Ĩ	e v	r v	с.	0 e	3	Ĩ		Ţ	344
CAA	TAT	ACC7	CAA	GCT(	- CGAC	CAT	- GGT	GCT	ATTO	- CATI	¥ TTC	TCI	raga	ATC	CTI		CAGA	AAT	CAT	L Ata	AAC	r GGA	G Atc	S AAC	A TGT	s 'ccr	N TCC	ם. ביתרים	341 1139
Î	Ĭ		s l	S	Т	M	V	Ļ	F	I	F	S	R	I	Ľ	n I	R	N	H	I	N	G	I	N	ç	Ğ	Ŝ	Ŝ	374
Q	Y V Doci		S	T	א תרבי	F	L	L	L	T	Y	A	N	Y	Ĺ	Ň	S	S	G	G	H	<u>а</u>	•	s	<u>ċ</u>	Ġ	T	T	370
H i	F 1		S	Q	I	R	G	F		K K	T	Q	V	E	Y Y	I I	L L	ĢGA	AAG K	aat Ņ	CCG P	ATG M	AAG K	ATG M	S	TAC Y	ATC M	SGTT V	1228 404
T V	v 1		ĸ	N	L	I	S	L	Å	ĸ	ĸ	å	v	D	¥	ł	ŗ	Ġ	Q :	N	P.	A	k :	I M	s	I Y	H M	v	400
GGA	FTTC F	GC A	GCA	AATZ Y	ACCC.	AAA K	GCA	ATT L	GCA1	CAC H	AGA	Geo	STCG	TCC	ATI T	:CC:	ITC	ATA	AAA K	GTT	CAC	CCA	GCG	AAG	GTG	GCT	ŢGC	TAAT	1318
ļ		; E	R	Ĩ	Ī	0	- H	v	Ĩ	Ī	Î	Ĭ	Ĩ	Ĩ	- т.	Ĩ,	Ĩ	v		Į.	<u>ו</u>	<u>ן</u> ו	N .	q	т т	ы Б	ĭ	Ĩ.	420
GCT	GT	TCI	CAG	- ATT <i>i</i>	ATTA	CAN	TTC	TGC	TAAT		AAT	ccc	- CAAT	ACC	- CAI	- 'GT(	GGGI	GCC	¥ ATT(	GTT	GGA	GGC	CCG	ATT	CAA	ATG	<u> </u>	ACG	1405
		. S	D	Ť	Y	N	s I	A	N	P	N	P	N I	T	H	Y	G	A I	I	V	Ģ	Ģ	P	I	Q	M	•	T	463
A (	5 E 7376	י. מידר בי	2 (77)	Y CDTC	L	Y CTDN	S	S	₽ ~~~~	P	N COT	P	N	I	L	v	G	À	İ :	L	Ġ	Ġ.	₽ >===	D	N	r 	D	S	459
F I	Ĩ		R	Ŝ	D	Ŷ	Ŝ	H	A	E	P	T	Ţ	Y Y	Ĩ	N N	A	A I	F	V I	A a	S	I .	S S	A SCT	L L	L		493
ŕ s	s r	D	Ŕ	N	N	Ý	Q	Q	S	Ė	Þ	A	Ŧ	¥	ł	ท่	Å	P :	L'	Ϋ́ν	G i	<b>a</b> :	L	A	F	F	A	Å	489
AAAACCTAATTAAACCAAATTCATGTAATAAATCCATAAGAAATTTCACACAACAAAGCCACTGCATGTGAACTACCTAC																													
NI	N P V T E 494																												
AATI	ANTANTIGCANGGTATTTTAGGTTCANATCACCCCATTGTTGTACTATTGTCTTTCATTAACAGTATAAATGGAAACTACACTTCTATG 1675												1675																
GGC1	IGTI	TTA	CCA	AAAA	AAA	AAA	AA	17	00																				

**Figure 1.** Endonuclease restriction map and nucleotide sequence of a bean abscission cellulase cDNA, pBAC10, and optimal alignment of bean and avocado (24) deduced amino acid sequences. The bar at the top of the figure is the endonuclease restriction map of the bean clone, pBAC10, and the bars immediately below this reflect the lengths of the bean clones discussed in the text. Within the amino acid sequences, predicted cleavage sites for the putative signal peptides are indicated with asterisks (26). Identically matched amino acids are indicated by connecting solid lines. Aligned cysteine residues in the mature polypeptides are indicated by a double underline. A single unmatched cysteine in the mature avocado sequence at position 106 is unmarked. Potential sites for glycosylation (Asn-X-Ser/Thr) in the avocado amino acid sequence are indicated by a single solid underline; no glycosylation consensus sequence is found in bean.

Amino acids co respective proteins	mprising the pre (see Fig. 1) have r	dicted sig lot been in	nal peptid cluded in th	es of thei nis analysis								
	A	Amino Acid Compositions										
Residue	Be	an	Avocado									
	number	mol %	number	mol %								
А	44	9.3	40	8.5								
С	6	1.2	7	1.4								
D	25	5.2	30	6.3								
E	17	3.6	16	3.4								
F	21	4.4	18	3.8								
G	37	7.8	38	8.1								
н	11	2.3	12	2.5								

Table I. Comparison of Deduced Amino Acid Compositions, Mol Wt, and pl of Bean Abscission and Avocado Fruit Cellulases ir

Residue	Be	an	Avocado				
	number	mol %	number	mol %			
Α	44	9.3	40	8.5			
С	6	1.2	7	1.4			
D	25	5.2	30	6.3			
E	17	3.6	16	3.4			
F	21	4.4	18	3.8			
G	37	7.8	38	8.1			
н	11	2.3	12	2.5			
I	23	4.8	14	2.9			
к	28	5.9	16	3.4			
L	34	7.2	45	9.5			
М	12	2.5	7	1.4			
Ν	24	5.0	25	5.3			
Р	19	4.0	22	4.6			
Q	16	3.3	20	4.2			
R	15	3.1	16	3.4			
S	58	12.2	52	11.0			
т	27	5.7	28	5.9			
v	20	4.2	26	5.5			
w	9	1.9	7	1.4			
Y	26	5.5	30	6.3			
D + E (acidic)	42	8.8	46	9.8			
H + K + R (basic)	54	11.4	44	9.3			
Residues	4	472	469				
Mol wt	51,8	317	51,395				
pl deduced sequence	7	7.8	5.3				
pl native protein		9.5	4	.7			

acteristics suggest that the two cellulases have similar secondary and tertiary structures. However, the pI values for the two cellulases are quite dissimilar. Native avocado fruit cellulase has a pI of 4.7 (1) and bean abscission cellulase a pI of 9.5 (8). The pI values deduced from their primary sequence, a measure that does not account for tertiary structure of the mature protein, are less disparate but still reflect the acidic and basic nature of the mature proteins, 5.3 and 7.8 for avocado and bean, respectively. The differences in pI values are primarily a result of 75% more lysine and 17% less aspartic acid in bean than in avocado cellulase (Table I).

Another notable difference between the two cellulases is that the mature avocado cellulase is a glycoprotein (3), whereas the mature bean cellulase does not appear to be glycosylated (6). The avocado sequence contains two consensus sequences for potential glycosylation (Asn-X-Ser/Thr) (22), whereas the bean cellulase does not contain this consensus sequence (Fig. 1).

#### Analysis of Genomic Southern Blot Data

A genomic Southern blot of HindIII and EcoRI digested DNA from avocado and bean was hybridized independently with both the avocado cellulase cDNA, pAV363, and the bean cellulase cDNA, pBAC10 (Fig. 3). The same blot was used for both hybridizations and the hybridization and washing conditions were designed to be approximately 40°C below Tm. When the avocado cellulase cDNA probe was hybridized to avocado genomic DNA, multiple restriction fragments were detected, suggesting the existence of a small gene family. When the same genomic blot was hybridized and washed at higher stringencies, approximately 15°C below Tm, the relative hybridization signal for several restriction fragments was diminished, further supporting the interpretation of a small gene family (24).

When the bean cellulase cDNA probe was hybridized to bean genomic DNA, a single HindIII fragment and two EcoRI fragments were detected. The bean cDNA clone has two EcoRI sites 48 nucleotides apart near the middle of the cDNA. The very small internal EcoRI fragment might not be detected on the Southern blot, but the presence of these sites in the middle of the cDNA predicts at least two larger restriction fragments for EcoRI digested bean genomic DNA. Thus, a single HindIII and two EcoRI bands suggest that the bean abscission cellulase is encoded by a single gene or a few very closely related genes. Moreover, the banding pattern detected was unchanged at higher stringency conditions (data not shown).

When the avocado cellulase cDNA probe was hybridized to bean genomic DNA, two restriction fragments in each lane of HindIII and EcoRI digested DNA hybridized to the avocado cDNA (Fig. 3). These restriction fragments are not the same restriction fragments that were detected with the bean abscission cellulase cDNA. Neither the avocado clone, pAV363, nor the bean clone, pBAC10, cross-hybridized with the respective restriction fragment(s) for the bean abscission and avocado fruit cellulase genes at the stringency employed (Fig. 3). In this regard, it is worth noting that sequence alignment of the bean and avocado clones showed an average of 64% nucleotide sequence identity. An average 64% sequence identity was apparently not sufficient to produce a hybridization signal under these stringency conditions. More-



Figure 2. Kyte and Doolittle (13) hydropathy plots for the deduced amino acid sequences of bean abscission and avocado fruit cellulases. The curve is the average of residue specific hydrophobicity over a window of nine residues. Hydrophobic sequences are indicated by a positive value and hydrophilic sequences are negative. Residue position in the sequences are indicated by numbers at the top and bottom of the graph. Stippled regions overlay putative signal peptides that would be cleaved from the mature polypeptides.



**Figure 3.** Genomic Southern blot of *Hind*III (H3) and *Eco*RI (R1) digested avocado and bean DNA. The blot was first probed with labeled purified cDNA insert from the avocado fruit cellulase cDNA (pAV363), and then washed with alkali and probed again with labeled purified cDNA insert from the bean abscission cellulase cDNA (pBAC10). Hybridization stringency conditions were 20% formamide and 0.9 M NaCl ( $5 \times$  SSPE) at 42°C. The stringency conditions of the final washes were 0.018 M NaCl ( $0.1 \times$  SSPE) at 45°C. Numbers to the left or right of each lane indicate fragment lengths in kilobase pairs.

over, within the bean cDNA probe at position 502–531 is a 29 nucleotide sequence of 68% G/C content (Fig. 1) which shares 93% identity with the avocado sequence. This conserved 29 nucleotide sequence within the full-length labeled probes must not have been of sufficient length to result in a significant hybridization signal under these conditions. Taken all together, these results suggest that there is at least one other gene in bean that has greater sequence similarity with the avocado fruit cellulase than with the bean abscission cellulase. The tissue-specific and developmental regulation of the expression of this other bean gene are completely unknown at this time.

No significant hybridization was detected when the bean cellulase cDNA probe, pBAC10, was hybridized to avocado genomic DNA (Fig. 3); however, when slightly lower stringency conditions were used, a faint hybridization signal was noted for a 6.7 kb fragment in the EcoRI digested avocado genomic DNA (data not shown). Interestingly, a 6.7 kb avocado DNA fragment is one of the hybridization signals seen in this lane at low stringency but not high stringency hybridization with the avocado cDNA, pAV363 (24). It is possible, but not yet determined, that this 6.7 kb fragment is part of an avocado abscission cellulase gene.

## DISCUSSION

Comparison of the nucleotide and deduced amino acid sequences for an avocado fruit cellulase and a bean abscission cellulase shows that these two proteins are similar in many respects, including primary (Fig. 1), secondary, and tertiary structure. Similarities in secondary and tertiary structures are indicated by alignment of cysteine residues (Fig. 1) and symmetry in the hydropathy plots of the deduced amino acid sequences (Fig. 2). However, there are notable differences between these two cellulases. The avocado cellulase is a glycoprotein (3), whereas the mature bean protein appears not to be glycosylated (6). Moreover, the isoelectric points of the native proteins are quite different. The pI of native avocado fruit cellulase is 4.7 (1) and bean abscission cellulase 9.5 (8). How differences in the glycosylation and charge of these proteins affect their secretion, cell wall location, or enzyme activity is unknown.

Both cellulases are expressed in specific organs and tissues at particular stages of development. It is not known, however, if the avocado fruit cellulase gene is also expressed during abscission of avocado leaves or if another related gene is expressed in avocado leaf abscission. In addition, it is not known if the bean abscission cellulase gene is expressed in other organs or tissues in bean, e.g. root initiation, anther development, or dehiscence. In this context it is interesting that the avocado fruit cellulase cDNA hybridized quite strongly to different restriction fragments of bean genomic DNA than did the bean abscission cellulase cDNA (Fig. 3). This suggests that there may be another gene in bean coding for a protein more similar to the avocado cellulase than to the bean abscission cellulase. This gene is most likely not expressed during bean leaf abscission due to its absence in the abscission specific cDNA library, but may be expressed in a different organ and stage of development. The genomic blot data suggest that there may be subfamilies of cellulases which are expressed in different tissues and stages of development and may have slightly different functional properties.

In 1987, the GenBank and EMBL sequence libraries were searched for sequences similar to the avocado cDNA. Additionally, the avocado cellulase sequence was compared directly to published sequences for bacterial and fungal cellulases. No significant similarities were found at that time (24). The GenBank and EMBL nucleotide sequence libraries were searched again in 1990 with both the avocado and bean cDNA sequences and, as before, no other nucleotide sequences were found to have significant sequence similarities with either of these cellulases. However, when the combined GenBank and EMBL nucleotide sequence libraries (26322 sequences) were translated in all three forward reading frames, a comparison search with the bean cellulase polypeptide sequence resulted in a match with the avocado cellulase (score 1322, the highest score) and a Pseudomonas fluorescens subsp. cellulosa carboxymethylcellulase (cellulase) gene (score 144, the second highest score) (GenBank locus X12570) (10). Optimal alignment of the open reading frame of the Pseudomonas cellulase and the bean abscission cellulase showed 27% identically matched amino acids. Curiously, a similar search of the combined GenBank and EMBL libraries with the avocado cellulase polypeptide sequence resulted in a match with itself (score 2448) but no other significant score. Optimal alignment of the Pseudomonas cellulase and the avocado cellulase peptide sequences showed only 22% identically matched amino acids. Additionally, the bean abscission cellulase was independently aligned with 23 different glucanases, including both endo- and exo-(1,4)- $\beta$ -D-glucanases as well as (1,3)-(1,4)- $\beta$ -Dglucanases. No bacterial, fungal, or plant cellulase or glucanase other than the avocado and *Pseudomonas* cellulases gave significant amino acid sequence alignments.

The *Pseudomonas* cellulase gene encodes a 962 amino acid polypeptide of 102 kD(10). The active portion of this cellulase resides within 608 residues from the amino terminal end (10). Interestingly, the bean sequence aligns optimally with the *Pseudomonas* sequence between residues 108 and 604. However, no obvious domain was noted in the alignments of these sequences. Amino acid matches were spread over a large fraction of the bean cellulase sequence. Possibly, with the identification of other plant cellulase cDNA clones, we will discover underlying similarities among plant, bacterial, and fungal cellulases which will aid in determining functional domains.

## **Note Added in Proof**

Since acceptance of this article, three articles have been published that deserve note. Two report amino acid sequences for bacterial enzymes that have significant sequence similarity with the avocado fruit cellulase. The third is a review of known cellulases wherein the avocado fruit cellulase was classified into one of the six families of cellulases, the E family. The E family of cellulases includes the Pseudomonas fluorescens cellulase discussed above. The articles are: Jauris S, Rucknagel KP, Schwarz WH, Kratzsch P, Bronnenmeir K, Staudenbauer WL (1990) Sequence analysis of the Clostridium stercorarium celZ gene encoding a thermoactive cellulase (Avicelase I): identification of catalytic and cellulosebinding domains. Mol Gen Genet 223: 258-267; Berger E, Jones WA, Jones DT, Woods DR (1990) Sequencing and expression of a cellodextrinase (cedl) gene from Butyrivibrio fibrisolvens H17c cloned in Escherichia coli. Mol Gen Genet 223: 310-318; Beguin P (1990) Molecular biology of cellulose degradation. Annu Rev Microbiol 44: 219-248.

#### ACKNOWLEDGMENT

We would like to thank Susan Baird for skillful technical assistance.

#### LITERATURE CITED

- Awad M, Lewis LN (1980) Avocado cellulase: extraction and purification. J Food Sci 45: 1625–1628
- Babbit JK, Powers MJ, Patterson ME (1973) Effects of growthregulators on cellulase, polygalacturonase, respiration, color, and texture of ripening tomatoes. J Am Soc Hortic Sci 98: 77-81
- Bennett AB, Christoffersen RE (1986) Synthesis and processing of cellulase from ripening avocado fruit. Plant Physiol 81: 830– 835
- Byrne H, Christou NV, Verma DPS, Maclachlan GA (1975) Purification and characterization of two cellulases from auxintreated pea epicotyls. J Biol Chem 250: 1012-1018
- Christoffersen RE, Tucker ML, Laties GG (1984) Cellulase gene expression in ripening avocado fruit: the accumulation of cellulase mRNA and protein as demonstrated by cDNA hybridization and immunodetection. Plant Mol Biol 3: 385-391

- del Campillo E, Durbin M, Lewis LN (1988) Changes in two forms of membrane-associated cellulase during ethylene-induced abscission. Plant Physiol 88: 904–909
- Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12: 387–395
- Durbin ML, Sexton R, Lewis LN (1981) The use of immunological methods to study the activity of cellulase isozymes (β-1,4glucan 4-glucan hydrolase) in bean leaf abscission. Plant Cell Environ 4: 67-73
- 9. Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25: 263–269
- Hall J, Gilbert HJ (1988) The nucleotide sequence of a carboxymethylcellulase gene from *Pseudomonas fluorescens* subsp. *cellulosa*. Mol Gen Genet 213: 112-117
- Horton RF, Osborne DJ (1967) Senescence, abscission, and cellulase activity in *Phaseolus vulgaris*. Nature 214: 1086–1088
- 12. Jackson AO, Larkins BA (1976) Influence of ionic strength, pH, and chelation of divalent metals on isolation of polyribosomes from tobacco leaves. Plant Physiol 57: 5-10
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132
- Langridge J, Langridge P, Bergquist PL (1980) Extraction of nucleic acid from agarose gels. Anal Biochem 103: 264–267
- 15. Lewis LN, Linkins AE, O'Sullivan SA, Reid PD (1975) Two forms of cellulase in bean plants. Proceedings of the Eighth International Conference on Plant Growth Substances. Hirokawa Publishing Co Inc, Tokyo, pp 708–718
- Linkins AE, Lewis LN, Palmer RL (1973) Hormonally induced changes in the stem and petiole anatomy and cellulase enzyme patterns in *Phaseolus vulgaris* L. Plant Physiol 52: 554–560
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi RG, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
- Sexton R, Roberts JA (1982) Cell biology of abscission. Annu Rev Plant Physiol 33: 133-162
- 21. Sexton R, del Campillo E, Duncan D, Lewis LN (1990) The purification of an anther cellulase ( $\beta$ (1-4)4-glucan hydrolase) from *Lathyrus odoratus* L. and its relationship to the similar enzyme found in abscission zones. Plant Sci **67**: 169–176
- Smith JA, Rottmann WL, Rubenstein L (1985) A study of the sites of glycosylation of zein proteins using the lectin concanavalin A. Plant Sci 38: 93-98
- Tucker ML, Laties GG (1984) Interrelationship of gene expression, polysome prevalence, and respiration during ripening of ethylene and/or cyanide-treated avocado fruit. Plant Physiol 74: 307-315
- 24. Tucker ML, Durbin ML, Clegg MT, Lewis LN (1987) Avocado cellulase: nucleotide sequence of a putative full-length cDNA clone and evidence for a small gene family. Plant Mol Biol 9: 197-203
- Tucker ML, Sexton R, del Campillo E, Lewis LN (1988) Bean abscission: characterization of a cDNA clone and regulation of gene expression by ethylene and auxin. Plant Physiol 88: 1257– 1262
- Von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133: 17-21