Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 9245–9249, September 1995
Plant Biology

Molecular cloning and sequence analysis of expansins—a highly conserved, multigene family of proteins that mediate cell wall extension in plants

(cellulose/cellulose binding proteins/plant cell enlargement/pollen allergens)

TATYANA Y. SHCHERBAN*, JUN SHI*, DANIEL M. DURACHKO*, MARK J. GUILTINAN†, SIMON J. MCQUEEN-MASON*, MARK SHIEH*, AND DANIEL J. COSGROVE*‡

Departments of *Biology and †Horticulture, Pennsylvania State University, University Park, PA 16802

Communicated by Hans Kende, Michigan State University, East Lansing, MI, July 7, 1995

ABSTRACT **Expansins are unusual proteins discovered by virtue of their ability to mediate cell wall extension in plants. We identified cDNA clones for two cucumber expansins on the basis of peptide sequences of proteins purified from cucumber hypocotyls. The expansin cDNAs encode related proteins with signal peptides predicted to direct protein secretion to the cell wall. Northern blot analysis showed moderate transcript abundance in the growing region of the hypocotyl and no detectable transcripts in the nongrowing region. Rice and *Arabidopsis* expansin cDNAs were identified from collections of anonymous cDNAs (expressed sequence tags). Sequence comparisons indicate at least four distinct expansin cDNAs in rice and at least six in *Arabidopsis*. Expansins are highly conserved in size and sequence (60–87% amino acid sequence identity and 75–95% similarity between any pairwise comparison), and phylogenetic trees indicate that this multigene family formed before the evolutionary divergence of monocotyledons and dicotyledons. Sequence and motif analyses show no similarities to known functional domains that might account for expansin action on wall extension. A series of highly conserved tryptophans may function in expansin binding to cellulose or other glycans. The high conservation of this multigene family indicates that the mechanism by which expansins promote wall extension tolerates little variation in protein structure.**

The growing cell wall in vascular plants consists of a complex network of cellulose microfibrils glued together by noncovalent interactions with matrix polymers (1). As a result of high turgor pressure, the cell wall bears considerable tensile stress. Growing cells enlarge by "loosening" the wall to allow slippage of the cellulose microfibrils and the adhering matrix polysaccharides and proteins (2). This results in relaxation of wall stress and turgor pressure and, consequently, an uptake of water to enlarge the cell and expand the wall (3). Despite considerable attention, the biochemical nature of wall loosening has long eluded definition (1).

Recently, we identified a class of wall proteins, called expansins, which have the remarkable ability to promote the extension of isolated plant cell walls (4). When growing cells are disrupted, for example by a freeze/thaw cycle, the walls retain significant capacity to extend when placed under tension and in acidic pH. This capacity to extend is lost upon denaturation of wall proteins (5) and may be restored by addition of purified expansins (4). Expansins appear to loosen the wall via an unusual nonhydrolytic mechanism that disrupts the noncovalent adhesion of wall matrix polysaccharides to one another or to cellulose microfibrils (6). Consistent with this mechanism, expansins were found to weaken cellulosic paper, which derives its mechanical strength from hydrogen bonding between cellulose fibers. Because expansins have an acidic optimum, they are ideal candidates for the protein mediator of "acid growth" in plants. Acid growth refers to the propensity of plant cells to grow faster when the wall pH is reduced (7, 8). This property is also observed in isolated cell walls under tension and has been reported in a wide variety of plant species.

In this study, we used peptide sequences from purified expansins to identify cDNA clones encoding these proteins. Sequence analysis of expansins from cucumber, *Arabidopsis*, pea, and rice indicates that these are unique proteins comprising an ancient, highly conserved, multigene family.§

MATERIALS AND METHODS

Protein Purification and Peptide Sequencing. We purified two expansin proteins from cucumber hypocotyls as described (4). In brief, walls from the growing hypocotyl of etiolated cucumber seedlings (*Cucumis sativus* L., cv. Burpee Pickler) were collected and extracted for ionically bound proteins, which were precipitated with ammonium sulfate and fractionated by HPLC on C_3 and sulfopropyl columns. Two active fractions, S1 and S2, were separated by SDS/PAGE, electroblotted onto a poly(vinylidene difluoride) membrane, and sent to the Wistar Protein Microsequencing Laboratory, Philadelphia, PA, for total amino acid composition and automated sequencing of tryptic peptides and undigested protein (9).

RNA Extraction, PCR Amplification, and cDNA Library Screening. RNA was extracted from 3-cm apical hypocotyl sections from 4-day etiolated cucumber seedlings. The segments were frozen in liquid nitrogen and extracted according to published methods (10). Poly(A)$^+$ RNA was selected by standard procedures (11) and used for cDNA synthesis with reverse transcriptase (RT) (Gibco/BRL; type M-MLV). A cDNA library was constructed with the λ ZAP II vector, using the Stratagene ZAP cDNA synthesis kit.

Degenerate primers based on the N-terminal amino acid sequence from the cucumber S1 expansin were used to amplify a 95-bp cDNA by PCR (sense primer, 5′-GAYTAYGGNGGNTGGCA-3′; antisense primer, 5′-TANARRTTNCCRTANCC-3′). This product was used as template for nested amplification with the primer 5′-CAYGCNACNTTYTAYGG-3′ and the same antisense primer, resulting in a 71-bp product, which was cloned into the Invitrogen TA cloning vector. The 71-bp product was radiolabeled by PCR with [α-^{32}P]dCTP, purified by centrifugation through a Bio-Gel P-6 column, and used to screen the

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.

Abbreviations: RT, reverse transcriptase; EST, expressed sequence tag.

‡To whom reprint requests should be addressed.

§The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U0382, U30460, U30476, U30477, U30478, U30479, U30480, and U30481).

cucumber cDNA library. Immobilization of bacteriophage and hybridization were carried out as described (11). Bluescript SK phagemids containing positive inserts were rescued from λ ZAP II using Exassist helper phage (Stratagene).

For S2 expansin, we designed S2-specific degenerate primers; the sense primer 5′-TTYACNGCNWSNGGNTGGGC-3′ was based on the S2 N-terminal peptide; the antisense primer 5′-CCNCCNCKRTADATNC-3′ was based on two peptide sequences (S2-22 and S2-55), which we predicted to be contiguous on the basis of homology with the S1 sequence (see *Results and Discussion*). Amplification with these primers was followed with a nested PCR amplification with the same antisense primer and a different sense primer (5′-CCNCCNCKRTADATNC-3′), which was based on the most conserved nucleotide sequence in the alignments of cucumber S1 expansin with several *Arabidopsis* and rice expansin cDNAs (see *Results and Discussion*). The amplification yielded a product of the predicted size (≈300 bp), which was cloned and sequenced to confirm an S2-specific partial clone of 312 bp. This fragment was used to screen the cDNA library.

DNA Sequencing. Plasmid DNA containing cDNA inserts was prepared by the Terrific Broth method (12) and used for dideoxynucleotide sequencing with Sequenase (United States Biochemical). We used subcloning and primer walking strategies to obtain complete sequences.

***In Vitro* Translation and Western Blotting.** DNA template was prepared by PCR amplification of the cucumber S1 expansin cDNA. Sense primer included the T7 RNA polymerase promoter followed by ATG and the first seven codons of the mature S1 expansin protein (5′-TAATACGACTCACTATAGGGAGCCATGGACTACGGTGGCTGGCAGAGC-3′). The antisense primer corresponded to the sequence following the stop codon (5′-TCATAGCAGTGTGGCTG-3′). DNA was purified with Promega's Wizard minicolumns and translated *in vitro* using Promega's TnT coupled wheat germ extract system. S1 protein synthesis was assessed by [^{35}S]methionine incorporation (data not shown) and by Western blot analysis using rabbit polyclonal antibodies directed against the cucumber S1 expansin (13).

Northern Blot Analysis. RNA was extracted as described above from the apical (growing) and basal (nongrowing) hypocotyl regions of 5-day etiolated cucumber seedlings. Fifteen micrograms of RNA was separated by electrophoresis in a 1.2% agarose gel with formaldehyde and Mops buffer and blotted onto Hybond-N membrane. S1 expansin cDNA was radiolabeled with [α-^{32}P]dCTP by random priming and hybridized to the filter (Amersham hybridization buffer tablets) at 65°C for 18 h and washed at the same temperature with 2× SSC for 15 min, then with 2× SSC/0.1% SDS for 15 min, and twice for 30 min. As a control, a blot was probed with a radiolabeled ubiquitin cDNA from *Arabidopsis thaliana*.

RESULTS AND DISCUSSION

Peptide Sequences. The following tryptic peptide sequences were obtained from purified S1 and S2 expansins: TGWQSMSR (S1-31), LTAYNLVPSNWQFGQTYEGPQ (S1-63), IGIYR (S2-22), GGIIPVLYQR (S2-55), VTTSDGQVQVFNNVVPSSXR (S2-59), and WGANWQSNSYLNGQSL (S2-64). The N-terminal sequence for S1 expansin was DYGGWQSGHATFYGGGDASGTMGGACGYGNLYSQ and that for S2 expansin was FTASGWAPAHATFYGESDASGXMGG. The N-terminal sequences showed that the two proteins were related, with 70% amino acid identity in the overlapping region, and were encoded by two distinct genes. Data base searches with these sequences failed to identify any closely related proteins.

S1 Expansin cDNA. A 71-bp cDNA was amplified by RT-PCR from cucumber mRNA using degenerate oligonucleotide primers designed from the N-terminal peptide sequence of the S1 expansin (see *Materials and Methods* for details). Using this cDNA as a hybridization probe, we identified 55 positive plaques from a λ ZAP II library of cucumber hypocotyl cDNA. Bluescript phagemids containing the *Eco*RI/*Xho* I cloned insert were rescued from 15 selected plaques and sequenced. All inserts contained the same sequence but of different lengths. Clone p332, containing the largest insert (1009 bp), was used for further analysis.

Three lines of evidence indicate that p332 contains an authentic cDNA for the cucumber S1 expansin. First, the protein encoded by the cDNA contained the peptide sequences from the purified S1 expansin protein (Fig. 1). Similarly, the amino acid composition and the general properties of the protein predicted from the cDNA corresponded closely to the characteristics of the purified S1 expansin (4)—e.g., size, pI, and presence of a signal peptide (see below). Second, *in vitro* translation of the cDNA produced a protein specifically recognized by antibody raised against the purified cucumber S1 expansin (Fig. 2*A*). Third, Northern analysis of mRNA from the growing and nongrowing regions of the cucumber hypocotyl (Fig. 2*B*) showed an expression pattern that matched the pattern of expansin activity found in this tissue—namely, high transcript abundance and expansin activity in the growing region of the hypocotyl and negligible transcript and activity in the nongrowing region (4). The Northern blotting results also indicate that expansin mRNA is moderately abundant in the growing hypocotyl because it can be detected in total RNA samples.

The cDNA sequence for the S1 expansin predicts a primary translation product of 250 aa and 27.2 kDa. The first 23 aa are recognized as a hydrophobic signal peptide by the (−3, −1) rule (14). Cleavage of the signal peptide would produce a mature protein with an N-terminal amino acid sequence matched exactly by that found in the purified protein. In plants, the presence of the signal peptide without other retention signals is sufficient to direct the protein to the wall via the endoplasmic reticulum–Golgi secretory pathway. This is consistent with previous findings that expansins are associated with plant cell walls (4). The resulting mature protein, after cleavage of the signal peptide, is predicted to consist of 227 residues (24.5 kDa). This size is ≈15% smaller than that

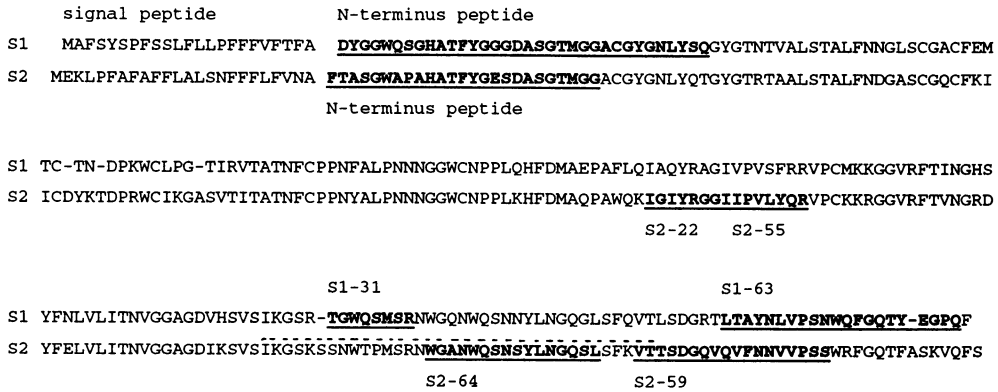

FIG. 1. Alignment of amino acid sequences predicted from cucumber S1 and S2 expansin cDNA clones. Underlined regions in boldface indicate the peptide sequences obtained from the purified proteins. A break is inserted between the predicted hydrophobic signal peptide and the start of the mature protein. Breaks in the alignment of the mature protein are indicated by dashes. Dotted line over S2 marks deleted region in the cDNA clone.

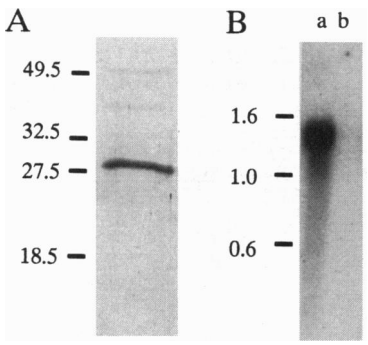

FIG. 2. (*A*) Western blot of *in vitro* translation product for the S1 expansin cDNA. Signal was dependent on the presence of cucumber S1 transcripts (control not shown). Size markers are in kDa. (*B*) Northern blot of total RNA from growing (lane a) and nongrowing (lane b) regions of the cucumber hypocotyl, detected with the S1 expansin probe. Size markers are in kb.

originally estimated from SDS/PAGE analysis of the purified protein (4). The predicted protein has a pI of 8.1, which resembles our previous estimate of 8.5 based on elution from ion-exchange columns (4). The amino acid composition of the S1 expansin is not remarkable, except that it is somewhat high in glycine (14.5%) and low in charged residues. Also, in comparison with average proteins (15), the S1 expansin is somewhat high in tryptophan (3%), cysteine (3.5%), and asparagine (9.25%). These properties are consistent with amino acid composition analysis of the purified protein (data not shown).

Analysis of the S1 expansin sequence by PROSITE (16) and BLOCKS (17) programs did not reveal any functional motifs that might account for the protein's biochemical activity—e.g., no consensus sequences for glucanohydrolase domains or cellulose binding domains. This is consistent with previous results indicating the protein acts by an unusual mechanism (4, 6, 18). The expansin protein does not contain repetitive domains, as judged by dot plots, and secondary structure predictions do not indicate any extraordinary features (data not shown).

S2 Expansin cDNA. We obtained a 312-bp partial clone of cucumber S2 expansin cDNA by RT-PCR using degenerate oligonucleotide primers based on the S2 expansin peptides (see *Materials and Methods*). This fragment was used to screen the cDNA library, resulting in a single isolate. Sequencing revealed that the clone apparently had intact 5′ and 3′ ends but contained a 109-bp deletion within the coding region, a deletion that resulted in a frameshift error and that omitted peptide S2-64 and part of peptide S2-59 (see Fig. 1). The sequence for the deleted region was obtained from cucumber hypocotyl mRNA by RT-PCR, using nested primers based on the region flanking the deletion and specific to the S2 sequence. A 337-bp PCR fragment containing the 109-bp deletion was amplified, cloned, and sequenced.

The predicted S2 expansin protein sequence is shown in Fig. 1. The primary translation product is 258 aa long (28.1 kDa). The (−3, −1) rule (14) predicts a signal peptide of 24 aa, cleavage of which should result in a mature protein of 234 aa (25.1 kDa) with an N-terminal peptide exactly matching that obtained from the purified protein. The mature protein has an overall amino acid identity to cucumber S1 expansin of 69%. As was the case with S1 expansin, no functional domains were identified by PROSITE and BLOCKS searches.

Expansin Homologs in *Arabidopsis*, Rice, and Pea Reveal a Highly Conserved Multigene Family. Using the S1 expansin amino acid sequence to search the nonredundant GenBank data base with the TBLASTN program (19), we identified a number of anonymous, partial cDNA sequences from the *Arabidopsis* and rice expressed sequence tag (EST) collections that appear to encode homologs of the cucumber expansins. These were identified by their high sequence similarity to the cucumber expansins (≈75% identity at the amino acid level). We identify the *Arabidopsis* expansins as AtEx1, AtEx2, etc., and the rice expansins as RiExA, RiExB, etc., in the order of their identification. The cucumber S1 and S2 expansins are referred to as CuExS1 and CuExS2. These names are not intended to imply specific homologies between individual expansin genes among these species.

In rice, there are at present 20 EST sequences with high similarity to cucumber expansin. They fall into five sequence classes, four of which encode distinct members of the expansin gene family: RiExA (GenBank accession no. D42004, with 3 redundant entries), RiExB (accession no. D24200, with 11 redundancies), RiExC (accession no. D47507), and RiExD (accession no. D23861). The fifth class (accession no. D40668 with 2 redundancies) appears to be a partial clone missing the N-terminal half of the coding region. Because the ESTs are only partially sequenced, it is not yet possible to determine whether the partial clone represents a fifth expansin in rice or overlaps with one of the other four expansins.

Upon sequencing RiExB and RiExD, we found the rice expansins to be remarkably similar to cucumber expansins (Fig. 3). For example, CuExS1 and RiExD have 79% amino acid identity, which is greater than that shared by the two cucumber expansins.

In a similar manner, we identified nine *Arabidopsis* ESTs that encode at least six distinct expansins, represented by the following names and GenBank accession nos.: AtEx1 (Z37585), AtEx2 (Z29931), AtEx5 (T45214), AtEx6 (T45828), AtEx8 (F14125), and AtEx9 (R29778). In addition to these six, there are currently three ESTs that appear to be partial cDNAs, missing the N-terminal region of the protein: AtEx3 (Z29178), AtEx4 (Z24541), and AtEx7 (T76481); their sequences reveal them to be distinct from each other, but it is not yet possible to determine whether they overlap the other ESTs. Four of the *Arabidopsis* ESTs were obtained and sequenced (Fig. 3); like the rice expansins, they closely resemble the cucumber sequences.

Thus, in rice there appear to be at least four distinct expansin genes and in *Arabidopsis* there are at least six. It should be noted that the rice ESTs come from a single cDNA library made from whole seedlings, whereas the *Arabidopsis* ESTs were obtained from multiple cDNA libraries made from several different plant stages and tissue cultures (but all of the Columbia ecotype). It seems likely that there are additional expansins still to be found in rice and *Arabidopsis*. The smaller number found in cucumber probably reflects the greater selectivity of the cDNA library (growing hypocotyls only).

Besides the anonymous cDNAs (ESTs) that appear to be expansins, the next closest group of sequences in the GenBank data base appear to be proteins identified as type 1 pollen allergens from several species (e.g., *Lolium* and *Phleum*, accession nos. M57474 and X78813). They show ≈25% amino acid sequence identity to expansins and are conserved in size and in several structural features of the expansin protein (see Fig. 3 and below). Statistical tests using the program RDF2 (20) confirm that the sequences for expansins and the type 1 pollen allergens are significantly related. Also in GenBank (accession no. X85187) is a sequence for a pea cDNA that is identified as a pollen allergen-like protein, but its sequence identifies it as an expansin homolog (Fig. 3).

The high conservation among the various expansins from rice (a monocotyledon) and cucumber, *Arabidopsis*, and pea (all dicotyledons) suggests several important points: (*i*) that the expansin multigene family arose by gene duplication before the evolutionary divergence of monocotyledons and dicotyledons some 150 million years ago; (*ii*) that the expansin protein has strict functional constraints that limit the structural modifications possible while maintaining function; and (*iii*) that the function of the protein is important to the normal development

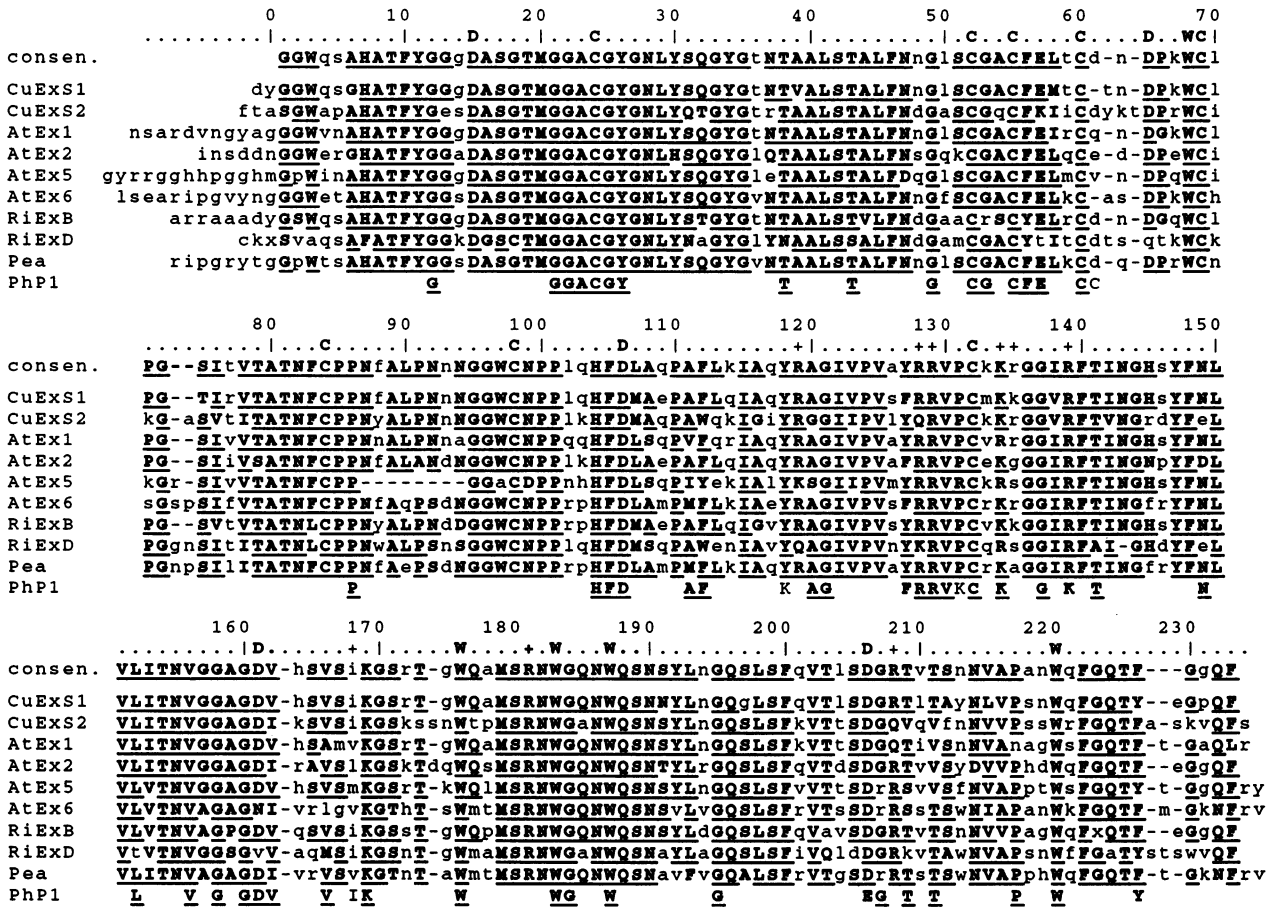

FIG. 3. Amino acid sequence alignments for nine expansin cDNAs from cucumber, rice, *Arabidopsis*, and pea. Also shown are aligned conserved residues from the *Phleum* type 1 pollen allergen (nonconserved amino acids are not shown). Residues identical to the consensus sequence in boldface are underlined, conservative substitutions are in boldface, and nonconservative substitutions are in lowercase. Numbering refers to the aligned consensus sequence, not individual sequences. Also shown on the number line are the conserved cysteines, tryptophans, and basic and acidic residues. Dashes indicate alignment gaps, and x indicates uncertain amino acid. Mature proteins, after signal peptide cleavage, are shown. Cleavage sites were predicted by the (−3, −1) rule with PCGENE software. The cleavage site for AtEx1 was not identified because the 5′ end of the cDNA is missing.

and physiology of angiosperms. These points are further supported by the fact that the signal peptide is not highly conserved among expansins; likewise, the 3′ and 5′ untranslated regions of expansin cDNAs are not conserved. Furthermore, the rate of nonsynonymous substitutions within the cDNA coding region for the mature protein is less than the rate of synonymous substitutions, indicating that negative (purifying) selection has worked against changes in the amino acid sequence of the protein.

To explore these points further, we constructed a phylogenetic tree (Fig. 4) based on nonsynonymous nucleotide substitutions using the MEGA program (21). The tree was rooted using the pollen allergen as outgroup. We note that the rice and cucumber expansins fall on either side of the root. This tree supports the idea that the duplication and functional specialization of expansin genes predated the divergence of monocotyledons and dicotyledons.

Phylogenetically Conserved Sites Suggest Functional Regions. Although PROSITE and BLOCKS analyses did not reveal how expansins might act, several phylogenetically conserved features of the expansin sequence merit comment. First, we note that the N-terminal half of the expansin protein contains a series of 8 conserved cysteines with spacing similar to that of conserved cysteines in the chitin-binding domain of wheat germ agglutinin (22). The cysteines in this lectin form a series of intramolecular disulfide bonds that stabilize the protein's structure. We suggest that the N-terminal half of the expansin protein may be folded and stabilized by disulfide bonds in an analogous manner.

Second, we note that the C-terminal one-third of the expansin protein contains four conserved tryptophans whose spacing resembles that of tryptophans in the cellulose-binding domains of cellulase (23, 24). Site-directed mutagenesis confirms the significance of tryptophan residues for cellulose binding (25). Moreover, tryptophan and the structurally re-

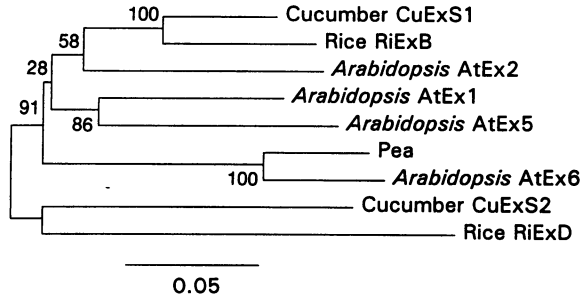

FIG. 4. Phylogenetic tree of the nine expansin sequences from Fig. 3. This tree was constructed by neighbor joining for nonsynonymous nucleotide differences using the MEGA program. Numbers indicate bootstrap *P* values. The tree was rooted with the PhP1 pollen allergen as outgroup. Similar results were obtained when trees were constructed using γ distances and amino acid *p* distances with Poisson correction.

lated amino acids phenylalanine and tyrosine are important for sugar binding in other proteins, and we note that these residues are particularly conserved in the C-terminal half of expansin. We therefore speculate that this region may be responsible for expansin binding to cellulose and related wall glycans (18).

Third, although charged residues are relatively underrepresented in the expansin protein, they are not randomly distributed along the protein backbone. The N-terminal half is nearly devoid of charged amino acids, whereas a short middle region (residues 119–139) contains the majority of the basic lysine and arginine residues (Fig. 3). These residues may be important for stabilizing protein structure by salt bridges or perhaps they interact with the carboxyl groups of wall pectins. Aspartic and glutamic acid residues may be particularly important for expansin function since the protein's acidic optimum (4) probably depends on protonation of the carboxyl groups of these residues. We point out in this regard that four glutamic acid residues (at positions 15, 106, 161, and 205; Fig. 3) are highly conserved among expansins.

Finally, we note that the conserved cysteines, tryptophans, and glutamic acids also appear as conserved residues between expansins and the pollen allergens (Fig. 3). This similarity suggests to us that the type 1 pollen allergen may function within the pollen cell wall as an expansin. This idea is supported by the fact that the N-terminal 23 aa of the *Phleum* PhP1 pollen allergen is recognized as a signal peptide by the (−3, −1) rule. Regarding the limited conservation between expansins and the type 1 pollen allergens, we note that the pollen spore coat and the pollen tube wall are composed of polymers different from those that make up the cell walls of the plant body and that the genes that function in the haploid and diploid phases of plant life may have diverged long ago.

Summary and Perspectives. Our previous studies of expansins and their mediation of acid growth in higher plants indicate that expansins promote sustained slippage between wall polymers without hydrolytic activity (4, 6, 18). We now have shown that expansins constitute an ancient multigene family that arose before the evolutionary divergence of monocotyledons and dicotyledons. If the mechanism of wall extension in land plants is monophyletic, as seems reasonable given the similarity in wall structure in these groups, we predict that expansins will be found in all groups of land plants. This prediction is strengthened by the observation of acid growth mechanisms in mosses (26), ferns (27), and gymnosperms (28) as well as in angiosperms.

At present, the function of the multiple forms of expansin (e.g., at least six in *Arabidopsis*) is unclear, but three possibilities occur to us. The various expansin genes may be expressed in different growing organs or in different cell types. The fact that we can find only two expansin proteins and corresponding cDNAs in the cucumber hypocotyl, yet there are at least six expansins in *Arabidopsis*, lends some support for this idea. We assume that cucumber has as many expansins as has *Arabidopsis*, but they are found in tissues other than the hypocotyl. An alternative idea (not mutually exclusive with the first) is that there are different forms of expansin that act on different components of the complex wall. This idea gains some support from the fact that the two cucumber expansins possess different biochemical properties (pH dependence, action on wall stress relaxation; see ref. 18). A variation on this idea is that two or more expansins may be expressed in the same cell but be regulated by different stimuli, such as hormones, light, and environmental stresses.

The novelty of expansin's sequence and its unique ability to induce plant cell wall extension present an unusual system for further study of how proteins may interact with polysaccharide networks to alter their rheological properties. By protein engineering and by analysis of more divergent expansins, it may be possible to identify the protein regions responsible for polysaccharide binding, pH dependence, and disruption of polysaccharide associations.

For technical support we thank Melva Perich and staff members at the Wistar Protein Microsequencing Facility. Dr. Blair Hedges assisted us with the phylogenetic analysis. For generous gifts of EST clones we thank Drs. Thierry Duprez (Institut National de la Recherche Agronomique, Versailles), Y. Nagamura and T. Sasaki (Rice Genome Research Program, Tsukuba), and Doreen Ware (Ohio State *Arabidopsis* Stock Center). This research was supported by Grant MCB-9317864 from the National Science Foundation to D.J.C.

1. Carpita, N. C. & Gibeaut, D. M. (1993) *Plant J.* **3,** 1–30.
2. Cosgrove, D. J. (1993) *Plant Physiol.* **102,** 1–6.
3. Cosgrove, D. J. (1993) *New Phytol.* **124,** 1–23.
4. McQueen-Mason, S., Durachko, D. M. & Cosgrove, D. J. (1992) *Plant Cell* **4,** 1425–1433.
5. Cosgrove, D. J. (1989) *Planta* **177,** 121–130.
6. McQueen-Mason, S. J. & Cosgrove, D. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 6574–6578.
7. Kutschera, U. & Schopfer, P. (1985) *Planta* **163,** 494–499.
8. Rayle, D. L. & Cleland, R. E. (1992) *Plant Physiol.* **99,** 1271–1274.
9. Reim, D. & Speicher, D. W. (1992) *Anal. Biochem.* **207,** 19–23.
10. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
11. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
12. Lee, S. & Rasheed, S. (1990) *BioTechniques* **9,** 676–679.
13. Li, Z.-C., Durachko, D. M. & Cosgrove, D. J. (1993) *Planta* **191,** 349–356.
14. von Heijne, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690.
15. McCaldon, P. & Argoes, P. (1988) *Proteins* **4,** 99–122.
16. Bairoch, A. (1992) *Nucleic Acid Res.* **20,** 2013–2018.
17. Henikoff, S. & Henikoff, J. G. (1994) *Genomics* **19,** 6565–6572.
18. McQueen-Mason, S. J. & Cosgrove, D. J. (1995) *Plant Physiol.* **107,** 87–100.
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
20. Pearson, W. R. (1990) *Methods Enzymol.* **183,** 63–98.
21. Kumar, S., Tamura, K. & Nei, M. (1993) MEGA: Molecular Evolutionary Genetics Analysis (Pennsylvania State Univ., University Park, PA), Version 1.01.
22. Wright, H. T., Sandrasegaram, G. & Wright, C. S. (1991) *J. Mol. Evol.* **33,** 283–294.
23. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. & Warren, R. A. J. (1991) *Microbiol. Rev.* **55,** 303–315.
24. Hamamota, T., Foong, F., Shoseyov, O. & Doi, R. H. (1992) *Mol. Gen. Genet.* **231,** 472–479.
25. Poole, D. M., Hazlewood, G. P., Huskisson, N. S., Virden, R. & Gilbert, H. J. (1993) *FEMS Microbiol. Lett.* **106,** 77–84.
26. Bittisnich, D. J. & Williamson, R. E. (1989) *Planta* **178,** 96–102.
27. Ridge, I. & Osborne, D. J. (1989) *Plant Cell Environ.* **12,** 383–393.
28. Lorences, E. P., Suelrez, L. & Zarra, I. (1989) *Plant Science* **62,** 53–61.