

Cloning and characterization of an 18 kiloDalton protein in the roots of the perennial weed *Taraxacum officinale* Weber (Dandelion) which has allergen- and pathogenesis-related protein properties

X.-Y. XU, J. D. BEWLEY & J. S. GREENWOOD

Department of Botany, University of Guelph, Guelph, Ontario N1G 2W1 Canada

ABSTRACT

The root of the persistent weed, dandelion (*Taraxacum officinale* Weber), contains a predominant 18 kDa protein which undergoes small seasonal fluctuations in amount, increasing in the late autumn months, and declining in the spring. This protein has been purified and found to consist of two major isoforms of pI 5.56 and pI 5.49. A full-length cDNA has been obtained, coding for the pI 5.56 isoform, and the 156-amino-acid sequence deduced. The protein shows homologies in amino acid composition to several allergen and intracellular pathogenesis-related proteins. The deduced protein does not contain a signal peptide nor any known organelle-targeting sequences, and thus is likely to be cytosolic. Expression of the 18 kDa protein gene is exclusive to the roots and stem tissues; transcripts accumulate during the late autumn months, and decline in the early spring. Changes in the amount of protein in the root are much less. The mRNA for the 18 kDa protein is not present in the dry seed, but appears in the roots within 16–18 h from the start of imbibition, and is expressed constitutively thereafter. Although it is the predominant protein in dandelion roots, its properties are different from those commonly associated with vegetative storage proteins.

Key-words: allergen; dandelion; pathogenesis-related protein; root protein; weeds.

INTRODUCTION

Perennial plants in temperate regions exhibit seasonal patterns of growth and development. Herbaceous perennial weeds undergo extensive growth during the spring and summer months, and then undergo extensive physiological changes in preparation for overwintering. The aerial parts of the plant senesce during the autumn months and there is a mobilization of nutrients to the persistent subterranean structures, which accumulate low-molecular-weight- and polymeric reserves (Woods, Harris & Caldwell 1957;

Heilmeier, Schulze & Whale 1986; Cyr, Bewley & Dunbroff 1990). In many roots carbohydrates are the primary source of reserve energy (Ho 1988; MacIsaac & Bewley 1995), although nitrogen-containing compounds are also a critical currency of storage metabolism (Suzuki & Kohno 1987; Cyr & Bewley 1989, 1990a,b; Hendershot & Volenec 1993).

The root of the perennial weed *Taraxacum officinale* L., the dandelion, contains a predominant 18 kDa protein throughout the year, which undergoes small fluctuations in amount, declining somewhat during the summer months and increasing in the autumn (Cyr & Bewley 1990b). Similar-sized polypeptides have been reported in the tubers of Jerusalem artichoke (*Helianthus tuberosus*) (Mussigman & Ledoigt 1989), and an 18 kDa protein is mobilized during sprouting. This latter protein from Jerusalem artichoke is antigenically related to the 18 kDa protein in dandelion roots; both species are members of the Compositae. Several low-molecular-mass proteins (15–32 kDa) are present in alfalfa (*Medicago sativa*) taproots (but not a similar 18 kDa protein); these accumulate during the autumn and early winter, and rapidly decline when spring growth commences (Hendershot & Volenec 1993; Cunningham & Volenec 1996). In white clover (*Trifolium repens*) 18 kDa proteins occur in the stolons, and two 15 kDa proteins in the roots are mobilized during early regrowth of defoliated plants (Corre *et al.* 1996). These observations have led to the suggestion that the predominant proteins in the roots and other subterranean tissues are vegetative storage proteins (VSPs), and are a temporary source of nitrogen, present in discrete storage vacuoles, which is mobilized and utilized when required for growth. The synthesis and degradation of VSPs in bark (Wetzel, Demmers & Greenwood 1989) and herbaceous perennials has been well-studied and aspects of their regulation are known (Staswick 1994). Many are seasonally regulated, and respond to changes in temperature and photoperiod (Clausen & Apel 1991; Coleman *et al.* 1991, 1992; Langheinrich & Tischner 1991).

The predominant 18 kDa protein in the dandelion roots is a small soluble polypeptide lacking structural complexity, and has been suggested to be a VSP (Cyr & Bewley 1990b). Little information is available on the protein itself. In this study we report on the purification and characteris-

Correspondence: Dr J. D. Bewley. Fax: (519) 767 1991; e-mail: dbewley@uoguelph.ca

tics of the 18 kDa protein, the isolation of its cDNA clone, and on the seasonal expression of its gene. A role for the 18 kDa protein as a VSP is questioned, however.

MATERIALS AND METHODS

Plant material

Dandelion (*Taraxacum officinale* Weber) roots were collected monthly on the campus of the University of Guelph from May 1993 and May 1995. The roots were copiously washed with tap water and distilled water, cut into small lengths (0.5 cm), frozen in liquid nitrogen and stored at -70°C until use. For protein purification, which was conducted in the Department of Plant Science, University of Adelaide, roots were collected from the Adelaide Hills during the winter months of June and July 1996 and washed and frozen as above.

Isoform determination

Lyophilized dandelion roots (0.5 g) collected in December were ground to a fine powder and extracted with 10 mL 0.1 M sodium phosphate buffer pH 6.8, with 10 mM cysteine and 3 mM phenylmethylsulphonyl fluoride (PMSF). Protein was precipitated by adding 4 vol. cold acetone and incubating overnight at -20°C . The pellet was air dried and dissolved in 9.5 M urea, 2% Nonident P-40, 0.8% each of Bio-Rad (Hercules, CA) ampholines pH 5–7, pH 6–8 and pH 3–10, and 5% β -mercaptoethanol. Two-dimensional electrophoresis [isoelectric focussing (IEF) followed by sodium dodecyl sulphate (SDS)–polyacrylamide gel electrophoresis (PAGE)] was according to O'Farrell (1975) before silver staining.

Protein purification

The protocol used was based on that of Hrmova & Fincher (1993). The optimal extraction buffer for the 18 kDa protein was determined to be 0.1 M Tris-HCl pH 7, 10 mM EDTA, 6 mM β -mercaptoethanol, 10 mM Na azide and 3 mM PMSF. Frozen stored roots (250 g) were extracted in 1 L chilled buffer using an homogenizer, stirred for 3 h at 4°C , centrifuged at 10 000 g for 30 min, and filtered through Miracloth (Fisher, Nepean, ON) to remove fine debris.

The 18 kDa protein was precipitated by 60–80% ammonium sulphate (Scopes 1994) at 4°C , and, following centrifugation at 10 000 g for 10 min, the pellet was dissolved in 100 mL DE52 column buffer (20 mM Tris-HCl pH 7, 6 mM β -mercaptoethanol). The dissolved protein was dialysed at 4°C in 8 L DE52 column buffer plus 10 mM sodium azide overnight at 4°C , being changed at least four times at 4 h intervals. The protein was then concentrated to a final volume of 40 mL in a 50 mL Amicon cell (Microcon, Amicon, Oakville, ON) using a 3 kDa cut-off membrane.

The concentrated protein solution was loaded onto an anion exchange DE52 column (Whatman, Fisher, Nepean, ON) and eluted with column buffer as a wash through.

When the UV absorption baseline of the eluate was stable, a linear gradient of NaCl (100 mL, 0 to 1 M NaCl in column buffer) was applied and 2 mL fractions taken. The fractions containing the 18 kDa protein were determined using SDS-PAGE and Coomassie Brilliant R-250 Blue staining (Fig. 2). These fractions (6–15) were pooled and dialysed against column buffer as for the DE52 column. Detection of the fractions enriched in 18 kDa protein (28–38, Fig. 2) was again by stained SDS-PAGE gels.

Fractions 28–38 were pooled and dialysed against chromatofocusing starting buffer (25 mM imidazole-HCl pH 7.4) and loaded onto a chromatofocusing column (pH 7–5, Polybuffer exchanger PBE94, Pharmacia, Baie d'Urfé, Quebec). After washing with starting buffer to establish a steady UV baseline the column was eluted with 1 : 6 diluted Polybuffer 74, pH 5 and two isoforms of the 18 kDa protein (Fig. 2) were detected on stained SDS-PAGE gels. Polybuffer was removed by repeating the BioGel P-60 (Bio-Rad) gel filtration.

The two purified 18 kDa isoforms were sequenced in the Department of Plant Science, University of Adelaide using a Hewlett-Packard (Camas, WA) G1005A protein sequencer and the HP 3.0 sequencing routine which is based on Edman degradation chemistry.

Rapid Amplification of CDNA ends-polymerase chain reaction

A BLAST search in Gen-Bank indicated that the 40-amino-acid sequence of the pI 5.56 isoform shared approx. 44–55% amino acid identity with Bet v1 from birch pollen and Mal d1 from apple leaves (Breiteneder *et al.* 1989; Vanek-Krebitz, Hoffmann-Sommergruber & Breiteneder 1995). A common motif (FKAF [V]) to the three proteins was used to make degenerate primers for Rapid Amplification of CDNA ends (RACE)-polymerase chain reaction (PCR).

Poly A⁺ RNA was prepared according to the extraction method of Pramanik *et al.* (1993) and oligo d(T)-cellulose affinity chromatography (Sambrook, Fritsch & Maniatis 1989) from 5 g frozen root powder collected in December 1994. Reverse transcription of the poly A⁺ RNA was achieved using the RiBoClone cDNA synthesis systems AMV RT kit (Promega, Madison, WI). The 3'- and 5'-end RACE PCR was conducted according to Frohman (1990) and Harvey & Darlison (1991), respectively, with minor modifications. The degenerate primer producing the highest amounts and most specific products was used for large-scale RACE-PCR amplification. The PCR products were isolated using 1% agarose gels according to Davis, Dibner & Battey (1986).

The RACE-PCR products were ligated into pGEM-T Easy vector and transformation of *Escherichia coli* XL-1 blue was conducted using standard protocols (Promega). Selected colonies containing the RACE-PCR products were amplified, the inserts digested from prepared plasmids and sequencing conducted with forward and backward M13 primers using an automatic DNA-sequencer. Forward and

backward sequence data from four clones of 3'-end RACE-PCR products and five clones of 5'-end RACE-PCR products were compared to each other using Align Plus and Gene Runner (Gen-Bank, Bethesda, MD, USA), and corrected sequences were overlapped using Align Plus to obtain the full length cDNA sequence for the 18 kDa protein. Characteristics of the protein were determined by PC/Gene (Gen-Bank) and Gene Runner after the latter was used to deduce the amino acid sequence.

Genomic DNA extraction and PCR

Genomic DNA was extracted from 5 g frozen dandelion leaves according to Ausubel *et al.* (1994). Five pairs of primers corresponding to specific regions of the cDNA were designed and standard PCR protocols, purification of the PCR products and their sequencing carried out.

Northern blot analysis

Total RNA extraction was according to Jepson *et al.* (1991) using 2 g frozen roots, and 10 µg total RNA was added to each lane of a denaturing agarose gel for separation before blotting onto a nitrocellulose membrane (Gene Screen, DuPont, NEN, Boston, MA) and probing with a 600 bp cDNA obtained by 3'-end RACE-PCR randomly labelled with [α^{32} P] dATP or [α^{32} P] dATP plus [α^{32} P] CTP. The membranes were washed in 2 × SSC (saline sodium citrate) + 0.1% SDS at room temperature for 5 min, 65 °C for 30 min and 0.1 × SSC and 0.1% SDS at 65 °C for 30 min before exposure to X-OMAT X-ray film (Kodak, Rochester, NY).

RESULTS

18 kDa protein purification

It has been shown previously (Cyr & Bewley 1990b) using one-dimensional SDS-PAGE that an 18-kDa protein is predominantly present in dandelion roots throughout the year, but in increased amounts from October to March. This protein runs as a single band in the presence or absence of reducing agents (β -mercaptoethanol or cysteine) (unpublished results) and hence is a monomeric protein. Two-dimensional electrophoresis shows there are two major 18 kDa proteins (Fig. 1) with pI values around 5.5–6. Two other proteins (19 kDa and 16 kDa) are also present in the roots, but are less predominant, and their pI values are higher and lower, respectively, than those of the 18 kDa protein.

Purification of the protein by chromatofocusing chromatography following ammonium sulphate precipitation, anion exchange DE52 column purification and gel filtration (Bio-Gel P-60) chromatography (Fig. 2B,C) yielded two peaks of 18 kDa protein with pI values of 5.56 and 5.49 (Fig. 2A). The peak fractions for each pI value were pooled separately and purified to near homogeneity on a second Bio-Gel P-60 gel filtration column.

A 40-amino-acid sequence at the N-terminal end of the pI 5.56 polypeptide was obtained, but the N-terminal end of the pI 5.49 polypeptide was blocked and could not be sequenced. A BLAST search in Gen Bank indicated that the 40-amino-acid sequence shared an approximately 44–55% amino acid identity with allergens Bet v1 from birch pollen and Mal d 1 from apple leaves (Breiteneder *et al.* 1989; Vanek-Krebitz *et al.* 1995), respectively, and

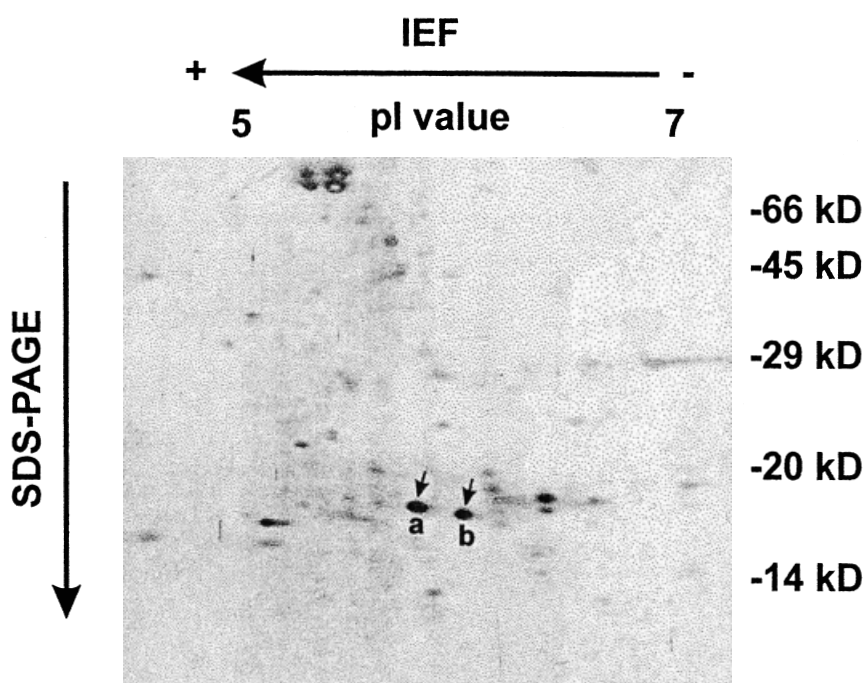


Figure 1. Two dimensional electrophoresis (IEF and SDS-PAGE) of dandelion root proteins. Total protein was extracted from roots collected in December in Guelph. Eighty micrograms of protein were loaded and after electrophoresis the gel was silver-stained. The arrows indicate the two major 18 kDa polypeptides.

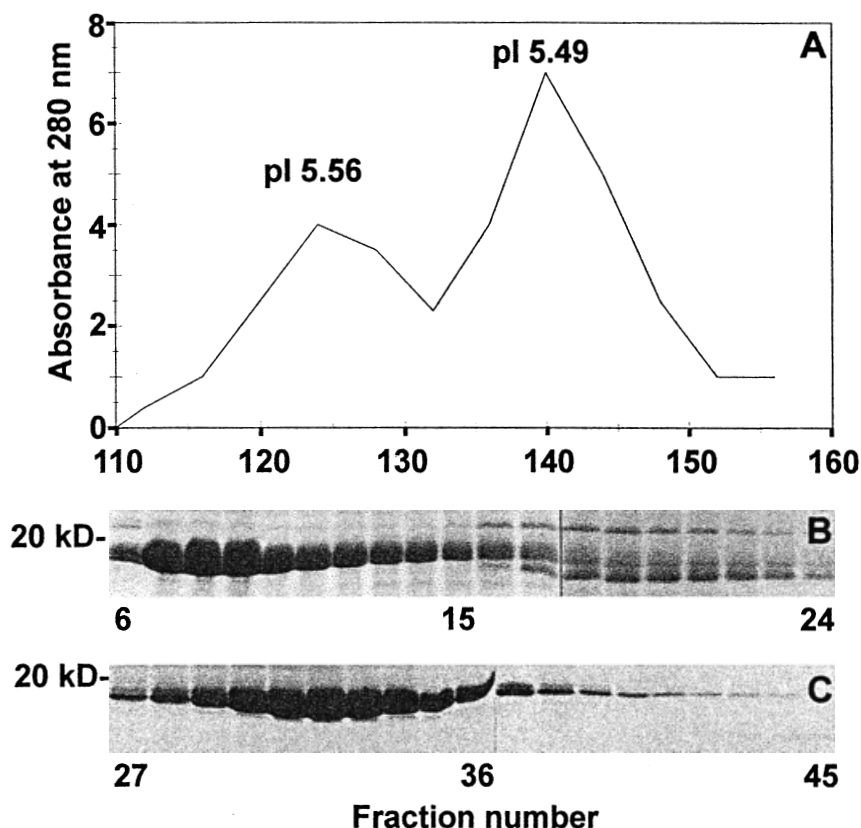


Figure 2. (A) Absorption at 280 nm of the eluate from a chromatofocusing chromatography column, showing the two isoforms of the 18 kDa protein. SDS-PAGE separation and Coomassie Brilliant Blue R-250 staining of the major fractions, 6–15 and 27–38, respectively, of the eluents from (B) an anion-exchange DE52 column and (C) a gel filtration BioGel P-60 column, showing the presence of the 18 kDa protein.

several intracellular pathogenesis related (IPR) proteins (e.g. disease response resistance protein from *Pisum sativum*; PVPR2 from *Phaseolus vulgaris*; STH-21 from *Solanum tuberosum*) (Fig. 3). However, even under mild stringency conditions there was no hybridization signal

obtained between mRNA extracts from dandelion root and cDNAs for either Bet v 1 or Mal d1.

A monoclonal antibody raised against purified recombinant Bet v 1 did not recognize the dandelion root 18 kDa proteins on western blots, whereas a polyclonal antibody

```

18 kD      ---AVAEFFETSSLSPSNIEKAEVIDFDTIARAEETYSIKTIEGDDGGVGTIK
PR1-3      MGVQKSEIVATSSVSAEKLEKGLCLDIITLLPRVLEGAIKSSETLEGDDGGVGTVK
pAPRI35    MGVFTYETETTSVIPPEKLEKAFILDAENLIEKVAETAVKGETILEGDDGGVGTIK
PRUA1      MGVFTYETETTSVIPPEKLEKAEVLDANLVKIAEQATKHEILEGDDGGVGTIK
Mal d1     MGVFNYYETETTSVIPAPRLKAFILDGNLIEKIAEQATKSTKIILEGDDGGVGTIK
Bet v1     MGVFNYYETETTSVIPAPRLKAFILDGNLIEKVAEQATSSVENIEGDDGGVGTIK

```

```

18 kD      SITYSDGVPTSSKHVDAIDSNMFSISYTIFFGDMGIISSGTHHLKFLPSA
PR1-3      LVHLGDASPFKTMKQVDAIDKATFTYSYIIDGIIILGFIESINNHTAVPNA
pAPRI35    KVTFGESQYAYVHRVIGLTKDNLSTYTLIEGDAISDVINIAVDILVASP
PRUA1      KITFGESQYGYVHKIISDKENYSYTLIEGDAISDVINIAVDILVASP
Mal d1     KVTFGESQYGYVQRVNGLTKDNLFTYSYSMIEGDTSDKLEKITYETKLIASP
Bet v1     KITFPEGSPFKYVKERVDEVHANFKYSYSMIEGDAISDVINIAVDILVASP

```

```

18 kD      DGGSVYKHSMVFKCKGDAKLTDENVSLMKGLKKTEKAIETVISHPEAC-
PR1-3      DGGCTVKTSTIIFNTKGDVVPENIKFANDQNLTIEKAVEAFLIAN----
pAPRI35    DGGSIIVKTTSHYHTKGDVEIKEQVKAGKEKAAGLEKLVEAYLLANPDAYN
PRUA1      SGGSIIVKTTSHYHTKGNVEIKEEHVKAGKEKASNLKLIETVYLGHPDAYN
Mal d1     DGGSIIVKTTSHYHTKGDVEIKEEHVKAGKEKASGLEKLVEAYLLAHSDAYN
Bet v1     DGGSIIVKTTSHYHTKGDHEMKAHEMKAIEKGEALLRAVESYLLAHSDAYN

```

Figure 3. Comparison of the amino acid sequence of the 18 kDa dandelion root protein with other proteins with similar sequences. PR1–3 is the pathogenesis-related protein from parsley; pAPRI35, PRUA1, Mal d 1 and Bet v1 are allergen proteins from apricot fruit, and cherry, apple and birch pollen. Amino acids which are the same or related are shaded and common motifs are underlined in bold.

detected the two pI isoforms of the protein, but weakly (not shown).

The 18 kDa protein clone and properties

The 3' and 5' RACE-PCR products were obtained using degenerate primers based on the 40-amino-acid sequence, and cloned into pGEM-T Easy vector (Promega). A 744 bp full-length cDNA sequence was obtained and the 156 amino acid sequence deduced (Gen Bank accession AF 036931) (Fig. 4). The N-terminal 40 amino acids were identical to those obtained by sequencing the purified protein.

Three amino acids are absent from the protein, Arg (R), Gln (Q) and Trp (W) (Table 1). The most abundant amino acids are Ser (S), Lys (K) and Ile (I), and charge density data obtained using Gene Runner show there are more negatively charged amino acids than positively charged ones (Fig. 5c). Gene Runner also indicates that the 18 kDa protein has a pI 5.6 (Fig. 5b), which is close to the pI 5.56 determined by chromatofocusing (Fig. 2). The predicted molecular mass of the protein is 16.911 kDa.

There are two potential N-glycosylation sites at residues

76 and 130, five protein kinase C phosphorylation sites at 36, 39, 51, 65, 142, and three casein kinase II phosphorylation sites at 56, 83 and 152.

The 18 kDa protein does not contain a signal peptide, as confirmed by the PC/Gene program which shows that no part of the protein exhibits the minimum hydrophobic value of a signal peptide (Fig. 5a). The PC/Gene program also indicates there is no eukaryotic secretory signal sequence, nor any known organelle-targeting sequences in the protein; the logical conclusion is that the 18 kDa protein is cytosolic. Light microscope sections of overwintering dandelion root have not revealed any protein-storing vacuoles (Butler 1996).

As expected from the homologies with the N-terminal 40-amino-acid sequence, the holoprotein shows high homologies with a group of allergen and IPR proteins (Fig. 3). The order of proteins demonstrating homology, from high to low, are: PR proteins PR1-3 and PR1-1 from parsley (64% similarity, 44% identity), pAPRI35 allergen proteins from apricot fruits (62% similarity, 42% identity), PRVA1 allergen protein from cherry (61%, 42%), Mal d 1 apple allergen protein (61%, 41%), Cor 1 apple allergen protein (59%, 42%), hornbeam allergen protein Car b 1 (63%, 39%), celery allergen protein API G1 (64%, 40%),

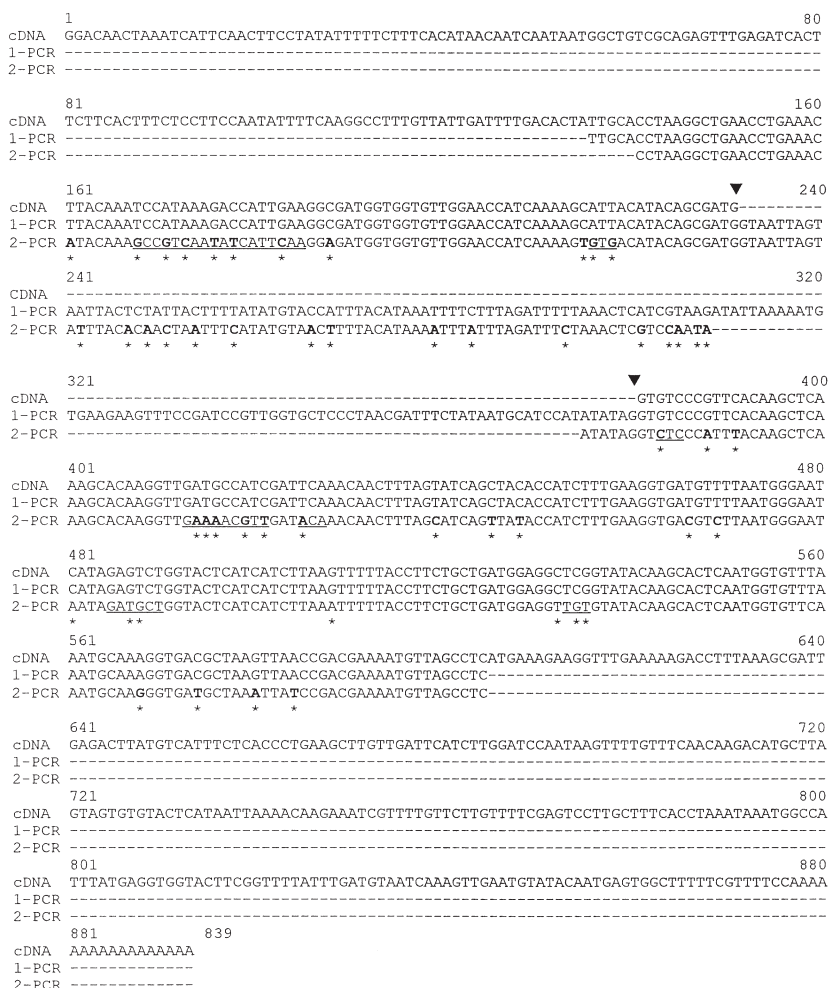


Figure 4. Comparison of the nucleotide sequences of genomic DNA-PCR products (1-PCR and 2-PCR) with the full-length 18 kDa protein cDNA sequence. Nucleotides different from those in the cDNA-sequence are in bold and asterisked, as are differences between 1-PCR and 2-PCR. Codons for different amino acids from those deduced from the cDNA sequence are underlined. Arrowheads show the start and stop sites of an intron. – indicates no matches between the sequences.

Table 1. Amino acid composition of the 18 kDa protein

Amino acid	Number of residues	Percentage by number
Ala (A)	10	6.41
Arg (R)	0	0
Asn (N)	4	2.56
Asp (D)	10	6.41
Cys (C)	2	1.28
Glu (E)	11	7.05
Gln (Q)	0	0
Gly (G)	12	7.69
His (H)	5	3.21
Ile (I)	15	9.62
Leu (L)	7	4.49
Lys (K)	16	10.26
Met (M)	3	1.92
Phe (F)	10	6.41
Pro (P)	6	3.85
Ser (S)	18	11.54
Thr (T)	12	7.69
Trp (W)	0	0
Tyr (Y)	5	3.21
Val (V)	10	6.41
acidic amino acids	26/156	16.67
basic amino acids	23/156	14.74
non-polar	73/156	46.89
polar	83/156	53.2

pathogen-induced allergen proteins from birch pollen Bet v 1-Sc 3 (59%, 41%) and Bet v 1 (61%, 38%).

Amino acid sequence alignments with five homologous proteins (Fig. 3) shows there are three blocks of amino acids in common; two are short (FKAF and DGGS) and one is long (EGDGGVGTIK). Sequence **GXGGXG** within the long common block is common to all proteins in this allergen and IPR group (Hoffmann-Sommergruber *et al.* 1997), and it seems the dandelion root 18 kDa protein is a member of this family.

A comparison of 18 kDa protein genomic and cDNA sequences

Two sequences were obtained from genomic DNA PCR products of 450 (1-PCR) and 400 bp (2-PCR) which corresponded to 18 kDa protein cDNA sequence obtained using 3' and 5'-end RACE amplification. The genomic DNA PCR products do not encode identical proteins, the sizes of their introns are different, and there are slight differences in the nucleotides within the intron and coding regions (Fig. 4). The introns and coding regions share 80.7 and 88% homology, respectively, at the nucleotide level, and an 85.7% identity in their deduced amino acid (105 residues) sequence.

The intron in the 450 bp product is 149 bp long, and in the 400 bp, 83 bp long (Fig. 4). Both introns are inserted in the same position, at amino acid 59 between nucleotide 1 and 2 (starting at nucleotide 232), which is the case for other allergen and PR proteins (Hoffman-Sommergruber *et al.* 1997) and are relatively AT rich (72.48% in 1-PCR and 79.52% in

2-PCR). A high AT content in the intron is characteristic of plant genes (Goodall & Filipowicz 1989), and the two exon-intron junctions are in accordance with the consensus plant splice (GT-AG) rule (Wolfe 1993). Moreover, all known genomic DNAs encoding allergen and IPR proteins contain only one intron, from 81 to 153 bp in length.

The sequence of the coding region of the 1-PCR product is identical to that of the cDNA obtained using 3'-and 5'-end RACE PCR amplifications, and therefore it is a partial gene sequence encoding the pI 5.56 18 kDa protein isoform. It is likely that the 2-PCR encodes the other isoform.

Southern blots using the cDNA encoding the 18 kDa protein showed three major hybridization bands with genomic DNA digested with EcoRI, Hind III and NcoI when the membranes were washed with $0.5 \times \text{SSC}$ at 65°C (Fig. 6A). This suggests there is a minimum of three gene copies encoding the 18 kDa protein. There were also some fainter hybridization bands, but analysis by Gene Runner of restriction sites indicated there are none present in the full-length cDNA and genomic DNA PCR products for the three restriction enzymes used. Hence the bands are probably a consequence of non-specific binding. Washing the membrane with $0.1 \times \text{SSC}$ at 65°C after hybridization resulted in only one band being present following each digest (Fig. 6B). Since there are two isoforms of the 18 kDa protein, with about 88% homology at the nucleotide level, it seems that the less strongly hybridized band between the probe and the gene encoding the pI 5.49 isoform was washed away under high stringency conditions. This is suggestive of the pI 5.56 isoform being encoded by one gene copy, and the other encoded by two.

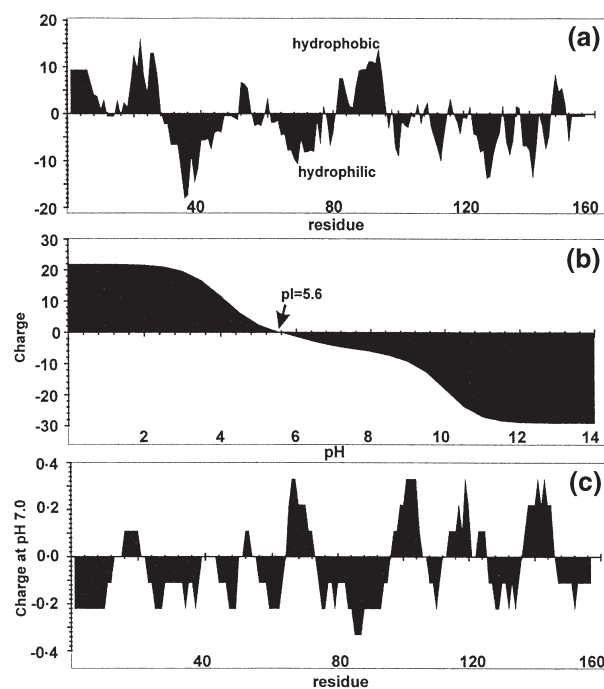


Figure 5. Properties of the 18 kDa protein showing its (a) hydropathy profile; (b) pI value; and (c) charge density.

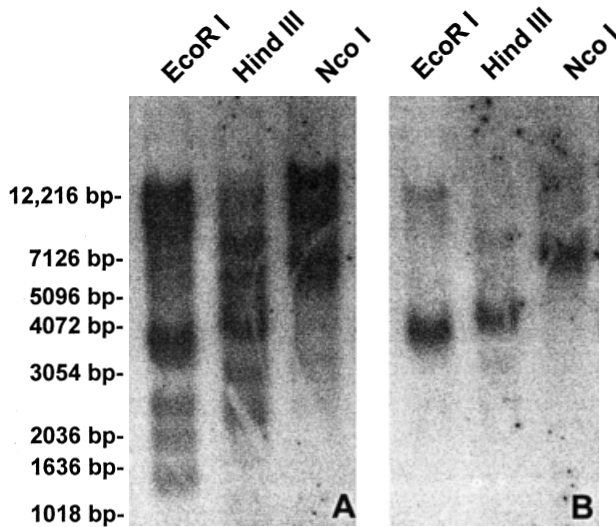


Figure 6. Southern blot of dandelion genomic DNA probed with the product of 3'-end RACE-PCR yielding part of the 18 kDa protein cDNA. (A) Membrane washed with $0.5 \times \text{SSC}$ at 65°C ; (B) membrane washed with $0.1 \times \text{SSC}$ at 65°C .

Expression of the 18 kDa protein gene in the root and stem

The mRNA for the 18 kDa protein undergoes seasonal changes in the dandelion root, with the least amount being present in May, a substantial increase occurring in September to November, and then a decline as spring approaches (Fig. 7). The seasonal fluctuations in the mRNA are much greater than those of the protein (Cyr & Bewley 1990a; Butler 1996; Xu *et al.* 2000). The expression of the gene

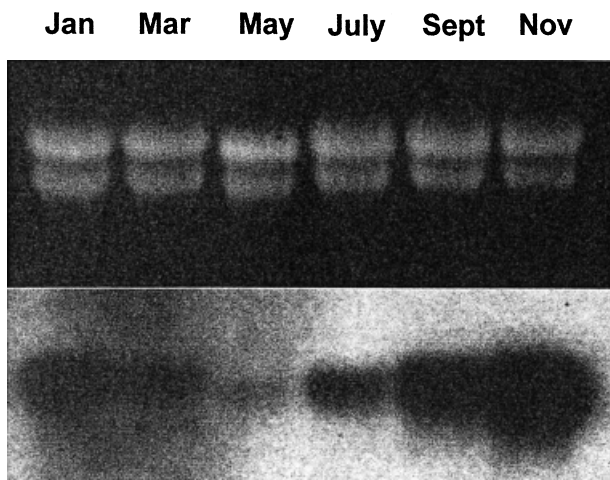


Figure 7. Seasonal regulation of expression of the gene encoding the 18 kDa protein. Top: Ethidium-bromide-stained RNA gel to show equal loading of RNA ($10 \mu\text{g}$ per lane). Bottom: Northern blot of RNA probed with radiolabelled cDNA encoding the pI 5.56 isoform of the 18 kDa protein.

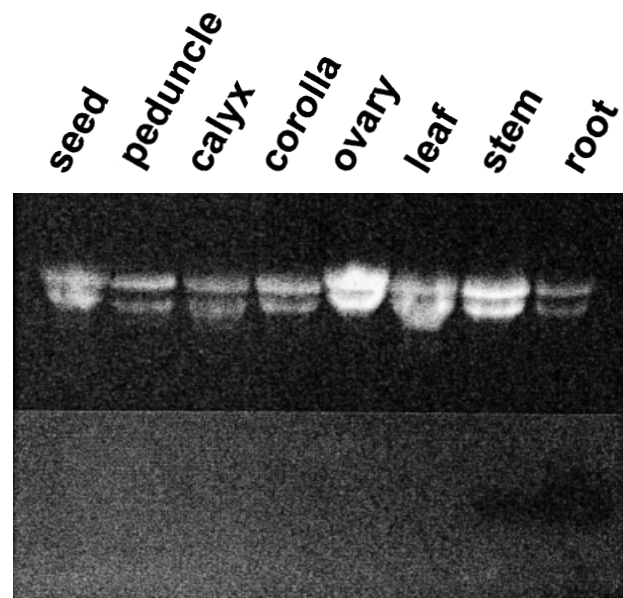


Figure 8. Spatial expression of the gene encoding the 18 kDa protein. Top and Bottom lanes as Fig. 7.

encoding the 18 kDa protein is specific to the roots and the stem (Fig. 8) the latter being a partially subterranean non-elongated tissue bearing the leaves of the rosette. Likewise, as expected, the 18 kDa protein is present only in the root and stem (Fig. 9); a low molecular mass protein is detectable in dry seeds, but this is 16–17 kDa in size and is probably not the same protein. There is no mRNA present in the dry seed for the 18 kDa protein (Figs 8 and 10).

The mRNA for the 18 kDa protein was present in seedlings from the first day, germination being completed

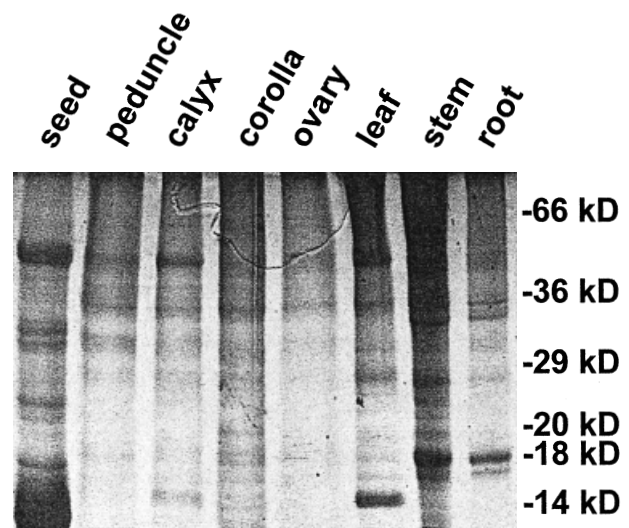


Figure 9. The presence of the 18 kDa protein in different parts of a dandelion plant. Equal amounts ($40 \mu\text{g}$) of protein were added per lane and following separation by SDS-PAGE they were stained with Coomassie Brilliant Blue R-250.

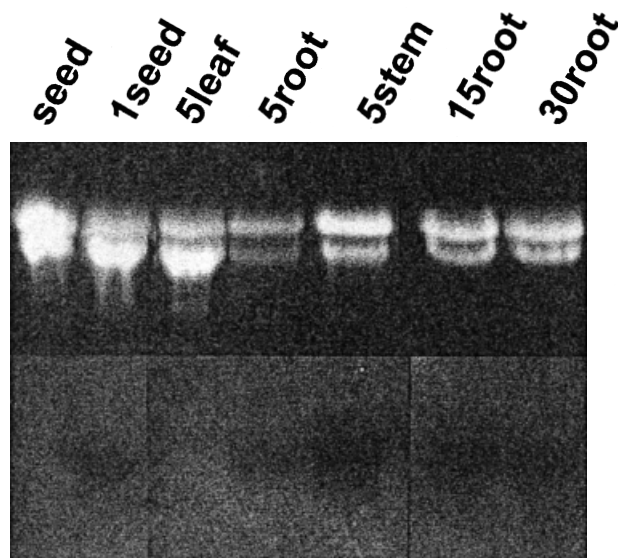


Figure 10. Expression of the gene encoding the 18 kDa protein during seedling establishment. Seed: dry seed. 1seed: 1-day-germinated seedling. 5leaf, 5root and 5stem: respective 5-day-old seedling parts. 15 and 30 root: 15- and 30-day-old roots. Top and Bottom as Fig. 7.

within 16–18 h from the start of imbibition. It was not present in 5-day-old leaf tissue (Fig. 10), nor in the older leaves (Fig. 8), but clearly in stem and root tissue after 5 d of growth. Thus the expression of the gene encoding the 18 kDa protein begins very early after germination in two specific regions of the dandelion plant and is expressed seasonally thereafter. SDS-PAGE analysis of extracted proteins shows a stainable 18 kDa band in 5- and 15-day-old stems and roots, but not in 5-day-old leaves or the 1-day-germinated seed (not shown). Either the amount of protein in the latter is too low for detection, or there is a lag period between transcription of the 18 kDa protein gene and translation of its mRNA.

DISCUSSION

The root of the perennial weed, dandelion contains a prominent 18 kDa protein, which increases a little with the approach of winter and declines during the spring (Cyr & Bewley 1990b; Xu 1998; Xu *et al.* 2000). The 18 kDa protein is composed of two major isoforms of pI 5.56 and 5.49 which were obtained after purification, and the N-terminal amino acid sequence of the pI 5.56 form obtained. Degenerate primers derived from this sequence were used for 3'- and 5'-end PCR RACE amplification and a 744-bp full-length cDNA and the deduced 156 amino acid sequence obtained for the pI 5.56 (pI 5.6 according to Gene Runner). The protein shows high homologies with several allergen and IPR proteins. A partial genomic sequence was obtained, as was part of the sequence of the other isoform, which is different by at least 11 amino acids.

Expression of the gene encoding the 18 kDa protein is

tissue specific, occurring only in the root and short stem. Expression occurs early following germination and the 18 kDa protein message appears to be constitutively expressed thereafter.

Dandelion, as a persistent perennial weed, inevitably encounters freezing stress while overwintering in its habitat in Guelph. The expression of the gene encoding the prominent 18 kDa protein is seasonally changeable, although fluctuations in the amount of protein do not reflect changes in the amount of message. During spring and summer the amount of specific mRNA is very low. Indeed, it is lower in March than in January, even though there is about the same amount of protein present (Cyr & Bewley 1990b). This might be due to a slow degradation of protein in the spring. During autumn and winter the protein increases only slightly between August and November, yet there is a large increase in its mRNA.

The role that the 18 kDa protein plays in the root is not known. Possibly it is a low-temperature-responsive (LTR) gene product with a role in cold protection, although its lack in variation in amount on a seasonal basis suggests otherwise. Unlike most LTR proteins, the 18 kDa protein does not show strong characteristics of dehydration-tolerance-related (DTR) proteins such as high hydrophobicity, with repeat motifs of serine, or lysine-rich repeats (Neven *et al.* 1993), but it is rich in these amino acids (11.54 and 10.26%, respectively). However, the LTR protein in alfalfa (pSM 2075) does not have serine clusters or lysine repeats either (Luo *et al.* 1991). The 18 kDa protein exhibits above 55% hydrophobicity and like many DTR proteins it lacks tryptophan. It also is likely to be a cytosolic protein in common with LTR proteins since it lacks a signal peptide and known organelle targeting motifs. It has been suggested (Houde *et al.* 1995) that the cryoprotective effect of the graminaceous WCS120 protein family is due to its abundance during cold acclimation. Given the enormous environmental range over which the dandelion grows, and the fact that in its elevated and northern-most climes exposure to cold temperatures could occur in any month of the year, the predominance of the protein even in summer might be part of its natural adaptation.

The 18 kDa protein also shares up to 44% homology with a group of allergen and IPR proteins, and it contains a motif common to all members of this group (Somssich *et al.* 1988). There are reports that some proteins which increase during cold acclimation also have similarities to PR proteins, e.g. in winter rye (Hon *et al.* 1995), and a positive correlation between cold acclimation and fungal resistance has been demonstrated in several freezing-tolerant grasses (Tronsmo 1984; Tronsmo *et al.* 1993) and osmotin in potato (Zhu, Chen & Li 1993). Decreasing the temperature of dandelion plants from 20 °C to 5 °C under controlled nitrogen conditions in a growth chamber did not result in a change in the amount of 18 kDa protein (Xu *et al.* 2000). Although we have no evidence that the 18 kDa protein plays either a role in freezing or cold-stress tolerance, or as a PR protein, its similarities with proteins involved in these processes are interesting.

A role for the 18 kDa protein as a vegetative storage

protein (VSP) seems even less likely. The synthesis and degradation of VSPs in bark, roots and tubers coincides strongly with seasonal changes (Coleman *et al.* 1991; Langheinrich & Tischner 1991; Staswick, Huang & Rhee 1991; Van Cleve & Apel 1993; Wetzel *et al.* 1989). In dandelion roots there is no major mobilization of the 18 kDa protein coincident with the rapid spring growth of the dandelion plant which would be expected to occur if the protein served as an important source of nitrogen. Moreover, the amount of protein remains high during the summer, when the mRNA available for replacement synthesis is low (i.e. high protein turnover in summer is unlikely), and when the most extensive growth and flowering takes place. Maintaining dandelion plants under greenhouse conditions for 9 weeks while feeding with 1, 5 or 25 mM NH_4NO_3 did not increase the amount of root storage protein (Butler 1996), indicating that its synthesis is not in response to nitrogen availability. This is unlike the situation in poplar, in which both the vegetative bark storage protein and its message increase substantially in response to as little as 2.5 mM NH_4NO_3 (Coleman, Bañados & Chen 1994). Taken together with the observation that the dandelion 18 kDa protein does not have the appropriate signals for deposition within the common VSP subcellular site, protein bodies, a role for this root protein as a VSP seems unlikely.

ACKNOWLEDGMENTS

This paper is dedicated by her coauthors and coadvisors to the memory of Dr Xiu-Ying Xu, who tragically died of cancer within weeks of completing her Ph.D. studies, of which these are a part. We are grateful to Professor G. B. Fincher, Department of Plant Science, University of Adelaide, who hosted Dr Xu in his laboratory during the protein purification experiments, and Mike Dalton and Dr Maria Hrmova for their expert advice. Dr Heimo Breiteneder and Dr Karen Hoffman-Sommergruber, University of Vienna, kindly provided antibodies and clones that aided in the isolation of the 18 kDa message.

REFERENCES

- Ausubel F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A. & Struhl K. (eds) (1994) *Preparation and analysis of DNA. Current Protocols in Molecular Biology. Section 2*. John Wiley & Sons, New York.
- Breiteneder H., Pettenburger K., Bito A., Valenta R., Krafr D., Rumpold H., Scheiner O. & Breitenbach M. (1989) The gene coding for the major birch pollen allergen Bet v I, is highly homologous to a pea disease resistance response gene. *EMBO Journal* **8**, 1935–1938.
- Butler S.M. (1996) An investigation of an 18 kD polypeptide in the subterranean organs of dandelion (*Taraxacum officinale* Weber). MSc Thesis, University of Guelph, Guelph, Canada.
- Clausen S. & Apel K. (1991) Seasonal changes in the concentration of the major storage protein and its mRNA in xylem ray cells of poplar trees. *Plant Molecular Biology* **17**, 669–678.
- Coleman G.D., Bañados M.P. & Chen T.H.H. (1994) Poplar bark storage protein and a related wound-induced gene are differentially induced by nitrogen. *Plant Physiology* **106**, 211–215.
- Coleman G.D., Chen T.H.H., Ernst S.G. & Fuchigami L. (1991) Photoperiod control of poplar bark storage protein accumulation. *Plant Physiology* **96**, 686–692.
- Coleman G.D., Chen T.H.H., Ernst S.G. & Fuchigami L.H. (1992) Complementary DNA cloning of poplar bark storage protein and control of its expression by photoperiod. *Plant Physiology* **98**, 687–693.
- Corre N., Bouchart V., Ourry A. & Boucaud J. (1996) Mobilization of nitrogen reserves during regrowth of defoliated *Trifolium repens* L. and identification of potential vegetative storage proteins. *Journal of Experimental Botany* **47**, 1111–1118.
- Cunningham S.M. & Volenec J.J. (1996) Purification and characterization of vegetative storage proteins from alfalfa (*Medicago sativa* L.) taproots. *Journal of Plant Physiology* **147**, 625–632.
- Cyr D.R. & Bewley J.D. (1989) Carbon and nitrogen reserves in the roots of pernicious perennial weeds. *Physiologia Plantarum* **77**, 67–72.
- Cyr D.R. & Bewley J.D. (1990a) Proteins in the roots of perennial weeds chicory (*Chicorium intybus* L.) and dandelion (*Taraxacum officinale* Weber) are associated with overwintering. *Planta* **182**, 370–374.
- Cyr D.R. & Bewley J.D. (1990b) Seasonal variation in nitrogen storage reserves in the roots of leafy spurge (*Euphorbia esula*) and responses to decapitation and defoliation. *Physiologia Plantarum* **78**, 361–366.
- Cyr D.R., Bewley J.D. & Dumbroff E.B. (1990) Seasonal dynamics of carbohydrate and nitrogenous components in the roots of perennial weeds. *Plant, Cell and Environment* **13**, 359–365.
- Davis J.G., Dibner M.D. & Battey J.F. (1986). *Methods in Molecular Biology*, pp. 123–125. Elsevier Science Publishing. The Netherlands.
- Frohman M.A. (1990) RACE: Rapid amplification of cDNA ends. In *PCR Protocols: a Guide to Methods and Amplification* (ed. M.A. Innis), pp. 28–38. Academic Press, San Diego.
- Goodall G.J. & Filipowicz W. (1989) The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell* **58**, 473–483.
- Harvey R.J. & Darlison G. (1991) Random-primed cDNA synthesis facilitates the isolation of multiple 5'-cDNA ends by RACE. *Nucleic Acids Research* **19**, 4002.
- Heilmeyer H., Schulze E.D. & Whale D.M. (1986) Carbon and nitrogen partitioning in the biennial monocarp *Arctium tomentosum* Mill. *Oecologia* **70**, 466–474.
- Hendershot K.L. & Volenec J.J. (1993) Taproot nitrogen accumulation and use in overwintering alfalfa (*Medicago sativa* L.). *Journal of Plant Physiology* **141**, 68–74.
- Ho L.C. (1988) Metabolism and compartmentation of imported sugars in sink organs in relation to sink strength. *Annual Reviews of Plant Physiology and Plant Molecular Biology* **39**, 355–378.
- Hoffmann-Sommergruber K., Vanek-Krebitz M., Radauer C., Wen J., Ferreira F., Scheiner O. & Breiteneder H. (1997) Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related protein share intron positions. *Gene* **197**, 91–100.
- Hon W.-C., Griffith M., Mlymarz A., Kowk Y.C. & Yang D.S.C. (1995) Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. *Plant Physiology* **109**, 879–889.
- Houde M., Daniel C., Lachapelle M., Allard F., Laliberte S. & Sarhan F. (1995) Immunolocalization of freezing-tolerance-associated proteins in the cytoplasm and nucleoplasm of wheat crown tissues. *Plant Journal* **8**, 583–593.
- Hrmova M. & Fincher G.B. (1993) Purification and properties of three (1 $\rightarrow$ 3)- β -D-glucanase isoenzymes from young leaves of barley (*Hordeum vulgare*). *Biochemical Journal* **289**, 453–461.
- Jepson I., Bray J., Jenkins G., Schuch W. & Edwards K. (1991) A

- rapid procedure for the construction of PCR cDNA libraries from small amounts of plant tissue. *Plant Molecular Biology* **9**, 131–138.
- Langheinrich U. & Tischner R. (1991) Vegetative storage proteins in poplar: induction and characterization of a 32- and a 36-kilodalton polypeptide. *Plant Physiology* **97**, 1017–1025.
- Luo M., Lin L., Hill R.D. & Mohapatra S.S. (1991) Primary structure of an environmental stress and abscisic acid-inducible alfalfa protein. *Plant Physiology* **17**, 1267–1269.
- MacIsaac S.A. & Bewley J.D. (1995) Seasonal changes in the starch content, and associated anabolic and catabolic enzymes, in the roots of the perennial weed, *Euphorbia esula*. *Plant Physiology and Biochemistry* **33**, 163–171.
- Mussigman C. & Ledoigt G. (1989) Major storage proteins in Jerusalem artichoke tubers. *Plant Physiology and Biochemistry* **27**, 81–86.
- Neven L.G., Haskell D.W., Hofig A., Li Q.-B. & Guy C.L. (1993) Characterization of a spinach gene responsive to low temperature and water stress. *Plant Molecular Biology* **21**, 291–305.
- O'Farrell P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry* **250**, 4007–4021.
- Pramanik S.K., Reynolds T.L., MacIsaac S.A. & Bewley J.D. (1993) Rapid and efficient purification of seed messenger RNA without phenol: chloroform extraction. *Seed Science Research* **3**, 137–139.
- Sambrook J., Fritsch E.F. & Maniatis T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Scopes R.K. (1994). *Protein Purification*, pp. 346–348. Springer-Verlag, New York.
- Somssich I.E., Schmelzer E., Kawalleck P. & Hahlbrock K. (1988) Gene structure *in situ* transcript localization of pathogenesis-related protein 1 in parsley. *Molecular and General Genetics* **213**, 93–98.
- Staswick P.E. (1994) Storage proteins of vegetative plant tissues. *Annual Review of Plant Physiology and Plant Molecular Biology* **45**, 303–322.
- Staswick P.E., Huang J.-F. & Rhee Y. (1991) Nitrogen and methyl jasmonate induction of soybean vegetative storage protein genes. *Plant Physiology* **96**, 130–136.
- Suzuki M. & Kohno K. (1987) Nitrogen levels and free amino acids in mulberry roots from autumn through spring. *Agricultural and Biological Chemistry* **51**, 53–58.
- Tronsmo A.M. (1984) Predisposing effects of low temperature on resistance to winter stress factors in grasses. *Acta Agriculturae Scandinavica* **31**, 210–220.
- Tronsmo A.M., Gregersen P., Hjeljord L., Sandal T., Bryngelsson T. & Collinge D.B. (1993) Cold-induced disease resistance. In *Mechanisms of Plant Defence Responses* (eds B. Fritig & M. Legrand). pp. 369–381. Kluwer Academic, The Netherlands.
- Van Cleve B. & Apel K. (1993) Induction by nitrogen and low temperature of storage protein synthesis in poplar trees exposed to long days. *Planta* **189**, 157–160.
- Vanek-Krebitz M., Hoffmann-Sommergruber K. & Breiteneder H. (1995) Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the major birch pollen. *Biochemical and Biophysical Research Communications* **214**, 538–551.
- Wetzel S., Demmers C. & Greenwood J.S. (1989) Seasonally fluctuating bark proteins are a potential form of nitrogen storage in three temperate hardwoods. *Planta* **178**, 275–281.
- Wolfe S.T. (1993) *Molecular and Cellular Biology*, pp. 592–594. Wadsworth, Belmont, CA.
- Woods F.W., Harris H.C. & Caldwell R.E. (1957) Monthly variations of carbohydrates and nitrogen in roots of sandhill oaks and wiregrass. *Ecology* **40**, 292–295.
- Xu X.Y. (1998) Physiology, biochemistry and molecular biology of a predominant 18 kDa protein in the roots of dandelion (*Taraxacum officinale* Weber). PhD Thesis, University of Guelph, Guelph, Canada.
- Xu X.Y., Butler S.M., Greenwood J.S. & Bewley J.D. (2000) Stress-mediated effects on the expression of a predominant 18 kDa *Taraxacum officinale* root protein. *Plant Physiology and Biochemistry* **38**, 491–497.
- Zhu B., Chen T.H.H. & Li P.H. (1993) Expression of an ABA-responsive osmotin-like gene during the induction of freezing tolerance in *Solanum emmersonii*. *Plant Molecular Biology* **21**, 729–735.

Received 9 April 2000; received in revised form 11 July 2000; accepted for publication 11 July 2000