PERGAMON

Phytochemistry 59 (2002) 279–291

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Serine proteinase inhibitors in the Compositae: distribution, polymorphism and properties

Alexander V. Konarev[a], Irina N. Anisimova[b], V.A. Gavrilova[b], T.E. Vachrusheva[b], G.Yu. Konechnaya[c], Mervyn Lewis[d], Peter R. Shewry[d,*]

[a]*All-Russian Institute of Plant Protection (VIZR), Podbelsky 3, St. Petersburg, 189620 Russia*
[b]*Vavilov Institute of Plant Industry (VIR), Bolshyaya Morskaya 44, St. Petersburg, 190000, Russia*
[c]*V.L. Komarov Botanical Institute, Prof. Popova 2, St. Petersburg, 197376, Russia*
[d]*IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, UK*

Received 11 October 2001; accepted 16 November 2001

Abstract

Multiple molecular forms of inhibitors of trypsin (TI) and chymotrypsin (CI), which are typical digestive enzymes of insects, mammals and micro-organisms, and subtilisin (SI), a proteinase of many bacteria and phytopathogenic fungi, were identified in seeds and vegetative organs of the majority of 128 wild and cultivated species representing 65 genera of three of the subfamilies of the Compositae. Inhibitors with M_r ranging from 7450 to 7800 and combining activities towards subtilisin and trypsin and/or chymotrypsin (T/C/SI) had the widest distribution and may be involved in plant defense mechanisms. They were found in many species of the subfamilies Carduoideae (genera *Carthamus*, *Centaurea*, *Cirsium*), Cichorioideae (*Lactuca*, *Taraxacum*) and Asteroideae (*Helianthus*, *Cosmos*, *Bidens*). Partial amino acid sequencing showed that the safflower (*Carthamus tinctorius*) T/C/SI and *Cosmos bipinnatus* T/C/SI, T/SI and C/SI belonged to the potato I inhibitor family. The most active, variable and heterogeneous inhibitors were found in species of the tribe Heliantheae, which is placed in the evolutionary advanced subfamily Asteroideae. Seeds of *Helianthus* species, *Eclipta prostrata*, *Gailardia aristata, Zinnia elegans* and *Silphium perfoliatum* contained various TI with M_r ranging from 1500 to 14,750, with some also containing SI. *H. annuus* seeds contain a unique cyclic TI of M_r 1514 and similar TI were also present in other *Helianthus* spp. and the related species *Tithonia diversifolia*. *Zinnia elegans* contained a TI with M_r 11,350 which appeared to represent a novel type of inhibitor distantly related to the cereal subgroup of Bowman–Birk inhibitors. TI and T/SI varied widely in *H. annuus* lines and wild *Helianthus* species in their presence or absence and composition. Similar T/SI components were found in the cultivated diploid *H. annuus* and annual diploid species with the B genome but not in perennials with the A genome. Some T/SI, SI and TI were detected in vegetative organs of sunflower and other Compositae. Studies of the polymorphism and distribution of proteinase inhibitors are relevant to the evolution of protective protein systems and the mechanisms of resistance to pathogenic organisms in the Compositae and other plants.

Keywords: Compositae; Proteinase inhibitors; Polymorphism

1. Introduction

The seeds and vegetative parts of plants contain various proteinaceous inhibitors of insect, fungal, mammalian and endogenous proteinases. These inhibitors may be involved in plant defense mechanisms against harmful organisms and may also play regulatory roles during plant development (Shewry and Lucas, 1997; Kumar et al., 1999). Furthermore, plant inhibitors are of interest in relation to host/parasite co-evolution (Konarev, 1996), as markers in studies of plant diversity and evolution (Konarev, 1982, 1996; Konarev et al., 1999a, 2002b; Kollipara and Hymowitz, 1992) and as potential drugs with antiviral and other properties. Genes encoding potent and stable inhibitors can also be transferred to other plants to improve their resistance to pests or fungi (Ryan, 1990). Proteinase inhibitors (PI) are well studied, particularly in the families Fabiaceae, Poaceae and Solanaceae. Some 12 families of inhibitors

* Corresponding author. Tel.: +44-1275-392181; fax: +44-1275-394299.

E-mail addresses: peter.shewry@bbsrc.ac.uk (P.R. Shewry), al_konarev@hotmail.com (A.V. Konarev).

0031-9422/02/$ - see front matter
PII: S0031-9422(01)00463-0

can be recognised based on their amino acid sequences and target proteinases (Shewry, 1999). The Compositae include important oilseeds (sunflower, safflower), vegetables (lettuce, artichoke, topinambour) and medicinal and even rubber plants (kok-saghyz). However, the inhibitors of seeds of the Compositae have not been studied until recently. Levitskii and Pogoreletskaya (1985) showed the presence of components which inhibited non-trypsin type proteinases of the pathogen *Botrytis cinerea* and inhibited fungal growth while Bhat et al. (1996) demonstrated the presence of inhibitors (presumably of trypsin) which inhibited the development of *Helicoverpa armigera* larvae. Similarly, the first characterization of a cysteine PI from sunflower seeds was only reported in 1996 (Kouzuma et al., 1996). We have previously isolated several isoforms of serine proteinase inhibitors from species of the Compositae using new approaches combining the identification of proteinase inhibitors after isoelectric focusing with purification by chromatography, including a unique cyclic trypsin inhibitor (TI) of M_r 1514 (Konarev et al., 1999a, 2000a; Luckett et al., 1999), trypsin/subtilisin inhibitors (T/SI) controlled by linked genes in sunflower (*Helianthus annuus* L.) seeds (Konarev et al., 1999a, 2000a) and various TI and T/SI in safflower and representatives of other taxa (Konarev et al., 2000b, 2002a). T/SI present in leaves and heads of sunflower were also shown to inhibit subtilisin-like extracellular proteinases of the white rot fungus *Sclerotinia sclerotiorum*, a major pathogen of sunflower, and proteinases of other fungi, indicating a possible protective role (Konarev et al., 1999b).

In the present study we aimed to (i) determine the distribution and polymorphism of proteinase inhibitors in cultivated and wild representatives of three main subfamilies of the Compositae: the Carduoideae, Cichorioideae and Asteroideae; (ii) purify and characterize novel inhibitor forms and (iii) determine the evolutionary relationships between inhibitors from various species of the Compositae and other plant taxa. Special attention has been given to inhibitors of trypsin and chymotrypsin which are typical digestive enzymes of insects, mammals and fungi, and subtilisin, a proteinase of phytopathogenic micro-organisms.

2. Results

2.1. Distribution of serine proteinase inhibitors in Compositae

The presence of serine proteinase inhibitors was demonstrated by isoelectric focusing followed by visualisation using the "gelatin replicas" method. In this method, the presence of inhibitors prevents the digestion of gelatin by various proteinases resulting in dark undigested "islands" on a transparent background. The results obtained for trypsin and subtilisin inhibitors in selected accessions are summarised in Fig. 1 and Table 1 (results for chymotrypsin inhibitors are not shown).

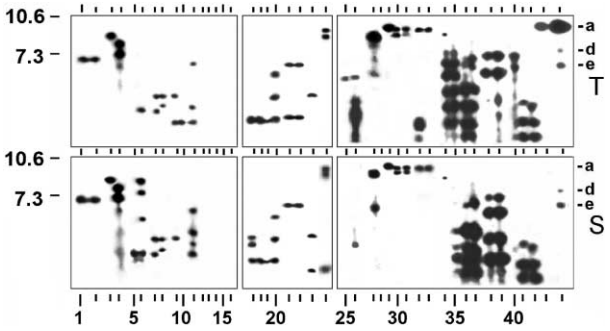

Fig. 1. Polymorphism of proteinase inhibitors in seeds and leaves of some representatives of Compositae (1–16, subfamily Carduoideae; 17–24, Cichorioideae; 25–44, Asteroideae). Proteins extracted with water were loaded on gels in volumes of 2–4 μl for slots 1–24 and 0.3–1 μl for slots 25–44 and separated by IEF in the pH range 3–10. Inhibitors were detected in two gelatin replicas made from each gel and developed with trypsin (T) and subtilisin (S). 7.3 and 10.6, positions of pI markers. a, d and e, positions of some *Helianthus annuus* inhibitor bands (Konarev et al., 2000b). **1** and **2**, *Carthamus tinctorius* and *C. oxyacanthus*. **3**–**6**, *Centaurea*: 3, *C. cyanus*; **4**, *C. triumfettii*; **5** and **6** *C. scabiosa* seeds and leaves. **7** and **8**, *Cirsium*: *C. arvense* and *C. vulgare*. **9** and **10**, *Carduus*: *C. acanthoides* and *C. crispus*. **11**, *Serratula coronata*. **12** and **13**, *Arctium*: *A. tomentosum* and *A. lappa*. **14**, *Cousinia badghysi*. **15**, *Saussurea salicifolia*. **16**, *Cynara scolymus*. **17**–**20**, *Taraxacum*: **17**–**19**, *T. officinale*, acc. of different origin; **20**, *T. hybernum*. **21** and **22**, *Lactuca*: *L. serriola* and *L. sativa*. **23**, *Cichorium intybus*. **24**, *Scorzonera hispanica* **25** and **26**, *Solidago virgaurea*, different acc. **27**–**30**, *Senecio*: **27**, *S. viscosus*, 28 and 29, *S. cineraria*, seed and leaf; **30**, *S. sakalavorum*, **31**, *Emilia sonchifolia*. **32** and **33**, *Rudbeckia*: *R. laciniata* and *R. hirta*; **34** and **35**, *Gaillardia aristata*, different acc. **36** and **37**, *Dahlia pinnata, acc*. **38** and **39**, *Cosmos bipinnatus*, acc. from St. Petersburg and Kenya. **40**, *Zinnia elegans*. **41** and **42**, *Bidens*: *B. tripartita* and *B. radiata*; **43**, *Tithonia diversifolia*. **44**, *H. annuus* VIR-104.

2.1.1. Subfamily Carduoideae

Seeds of 25 acc. of cultivated safflower (*C. tinctorius*) originating from various parts of the world and wild *Carthamus* L. species, representing species groups with different chromosome numbers, had a single component of pI about 7.0 which was inhibitory to trypsin, subtilisin (Fig. 1, T and S, tracks 1 and 2) and chymotrypsin (T/C/SI). However, species of *Centaurea* L., which is taxonomically close to *Carthamus*, showed variation in inhibitors. Seeds of *Centaurea cyanus* and *C. triumfettii* (Fig. 1, tracks 3 and 4) possessed active T/C/SI components with pI 7.3 and above while *C. scabiosa* (Fig. 1, track 5) and *C. jacea* (not shown) had only weak SI with pI of about 5.5. All accessions of *C. cyanus* had the same T/C/SI band with pI approx 9.0. Leaves of *C. cyanus* did not contain inhibitors of the serine proteinases studied, in contrast to *C. scabiosa* leaves (Fig. 1, track 6) which had one T/SI and several SI components.

Seeds of *Cirsium arvense* (Fig. 1, track 7) and *C. vulgare* (Fig. 1, track 8) differed in their compositions of

Table 1
Distribution of proteinase inhibitors in seeds of some species of the Compositae (based on data from IEF, TLGF and the gelatin replicas method)

Subfamily, **tribe**, subtribe, genus, species	T/(C)/SI	TI	S/(C)I	**Subfamily**, **tribe**, subtribe, genus, species	T/(C)/SI	TI	S/(C)I
Carduoideae	+/−	−	+/−	***Tageteae, Tagetes*** sp.	−/w?	−	−
Carthamus sp.	+	−	−	***Inuleae*** (2 genera)	+/−	−	−
C. tinctorius (25 acc.)	+	−	−	*Helichrysum bracteatum*	+	−	−
Centaurea sp.	+/−	−	+/−	*Inula helenium*	−	−	−
C. cyanus, *C. triumfetii*	+	−	−	***Heliantheae*** (20genera)	+/−	+/−	+/−
C. scabiosa	−	−	+	*Echinacea purpurea*	−	+w	+w
Cirsium sp.	+w	−	−	*Arnica iljinii*	+!	−	+
Carduus sp.	+w	+w	−	*Bidens* (3 sp.), *B. pilosa*	+!	−	−
Serratula coronata	+	+	+	*Coreopsis*	+	−	−
Arctium, Cousinia, Saussurea, Cynara sp.	−	−	−	*Cosmos bipinnatus*	+!	−	+!
Cichorioideae, ***Lactuceae***	+/−	+/−	+/−	*Cosmos caudatus*	+	−	+
Cichorium intybus	+	−	+	*Dahlia pinnata*	+!	−	−
Lactuca sp., *L. sativa* (23 acc.)	+	+	−	*Echinacea purpurea*	−	+w	+w
Scorzonera hispanica	+	−	+	*Eclipta prostrata*	−	+!	+w
Sonshus sp.	+w	−	−	*Galinsoga parviflora*	−	−	−
Taraxacum sp.	+	−	+/−	*Gaillardia aristata*	+	+!	−
T. officinale	+	−	+	*Rudbeckia laciniata*	+	+!	+
Tragopogon pratense	−	+	−	*Rudbeckia hirta*	+	−	−
Asteroideae	+/−	+/−	+/−	*Rudbeckia speciosa*	+	−	+
Astereae (6 genera)	+/−	+/−	−	*Silphium perfoliatum*	−	+!	−
Solidago virgaurea	+	+!	−	*Zinnia elegans*	+	+!	−
Erigeron uniflorus	−	+	−	*Helianthinae*	+/−	+/−	+/−
Anthemideae (8 genera)	−	−	−	*Helianthus* (18 sp.)	+/−	+H!/−	−
Chrysanthemum sp.	−	−	−	*H. annuus*	+	+H!/−	−
Senecioneae: *Emilia*, *Senecio* sp.	+	−	−	*H. tuberosus*	+w	H!	−
Senecio viscosus	+!	−	−	*Tithonia diversifolia*	+w	H!	−
Eupatorieae (2 genera)	+/−	−	−	*Simsia*, *Enceliopsis, Alvordia*,			
Eupatorium cannabinum	+w	−	−	*Florencia*, *Wedelia, Viguiera* sp.	−	−	−
Calenduleae: *Calendula officinalis*	+	−	−				

"+" or "−", Indicate presence or absence of inhibitors; "+/−", indicates variability inside taxon in the presence/absence of inhibitors; "!" and "w", indicate very active and weak IEF inhibitor bands; "H", indicates the presence of TI related to SFTI-1.

weak T/SI but no inhibitors were found in the leaves. Seeds of *Carduus* species had weak TI (Fig. 1, tracks 9 and 10) and T/SI (Fig. 1, track 10). *Serratula coronata* seeds (Fig. 1, track 11) contained a heterogeneous mixture of proteinase inhibitors but none were detected in seeds of burdock (*Arctium*), artichoke (*Cynara*), *Cousinia* or *Saussurea* species (Fig. 1, tracks 12–16).

2.1.2. Subfamily Cichorioideae, tribe Lactuceae

All accessions of *Lactuca* (23 var. of *L. sativa*, *L. serriola, L. livida* and *L. quercina*) had the same pattern of inhibitors with T/SI with pI about 7.3 and TI with low pI (about 5.0) as shown in Fig. 1 (tracks 21 and 22). *Taraxacum* species (Fig. 1, tracks 17–20) differed in their spectra of T/SI and SI, as did accessions of *T. officinale* (tracks 36–38). Seeds of some accessions of *Cichorium intybus* (Fig. 1, track 23) and *Scorzonera hispanica* (Fig. 1, track 24) contained T/SI and SI while species of other genera contained only weak inhibitors or lacked them.

2.1.3. Subfamily Asteroideae

Analysis of 46 genera from 8 tribes showed that the least active inhibitors were present in seeds of representatives of the tribes Anthemideae, Eupatorieae and Tageteae, with inhibitors not being detected in some such as *Chrysanthemum* and *Tagetes* species. The most active, variable and heterogeneous inhibitors were found in seeds of the tribe Heliantheae Cass. with TI, T/SI, T/C/SI and C/SI being detected (Fig. 1, tracks 32–44 and Table 1). Species of the tribes Asterea (*Solidago virgaurea* with T/SI and TI, Fig.1, tracks 25 and 26), Senecioneae (*Senecio* and *Emilia* spp. with T/SI, Fig. 1, tracks 27–31), Calenduleae (*Calendula* and *Dimorphotheca* spp. with T/SI, not shown) and Inuleae (*Helichrysum bracteatum*, T/SI, not shown) were intermediate in inhibitor activity. *Arnica iljinii* seeds contained potent T/SI, in contrast to *A. alpina* (not shown), while *Rudbeckia* species differed in the presence or absence of very active TI (Fig. 1, tracks 32 and 33) with pI near 5.0. *Gaillardia aristata* possessed relatively weak T/SI and potent heterogeneous TI while two *Dahlia pinnata* acc. (Fig. 1, tracks 36 and 37) had similarly heterogeneous and highly active T/C/SI. *Bidens* species had closely similar spectra of T/C/SI (Fig. 1, tracks 41 and 42). Two *Cosmos bipinnatus* acc. had similar spectra of inhibitors consisting of T/C/SI and C/SI (Fig. 1,

tracks 38 and 39) and inhibitors of the same type but with different pI were detected in *C. caudatus* (not shown). *Zinnia elegans* (Fig. 1, track 40) and *Silphium perfoliatum* (not shown) had potent TI. *Helianthus annuus* and *Tithonia diversifolia* had similar highly alkaline TI components (Fig. 1, tracks 43 and 44) but different T/SI. The T/SI of *T. diversifolia* are not visible in Fig. 1 due to the low sensitivity of the conditions used for detection. We failed to find inhibitors in the other representatives of subtribe Helianthinae listed in Table 1.

2.1.4. Genus Helianthus L.

Analysis of seeds of many varieties and lines of cultivated annual sunflower (*H. annuus*) revealed seven main band positions (Fig. 2, bands a–f and h) when gelatin replicas were developed by trypsin (T) and five when subtilisin was used (Fig. 2, bands c–f and h). Bands a–f and h were active against both proteinases (i.e. were T/SI) while a and b were TI (with band b being a doublet). Since the inhibitor components corresponding to band a were much more active than others they were analyzed separately. Band a corresponded to the major cyclic TI with M_r 1,514 (SFTI-1) described previously (Konarev et al., 1999a, 2000a; Luckett et al., 1999) and was present in *H. annuus*, other diploid annual species including *H. petiolaris*, *H. debilis*, *H. nuttallii*. and *H. praecox*, in all the tetraploid and hexaploid *Helianthus* species including topinambour (*H. tuberosus*) and in all three acc. of *Tithonia diversifolia* (not shown). Some *H. annuus* lines (VIR-369 and VIR-848b) and single seeds of one of two *H. hirsutus* accessions lacked band a which was also absent from *H. annuus* leaves and heads and from *H. tuberosus* L. leaves and tubers. TLGF under non-denaturating conditions showed that the native TI present in sunflower seeds and SFTI-1 purified by affinity chromatography and RP–HPLC (Fig. 3, tracks 11 and 14) had similar masses indicating the absence of significant modification of the protein during purification. It also confirmed the presence of similar low M_r inhibitors in the majority of other *Helianthus* species including *H. tuberosus* (Fig. 3, track 12) and in *T. diversifolia* (Fig. 3, track 13) but not in any other species of the Compositae.

T/SI are highly variable in *H. annuus* lines, in wild annual forms and in diploid, tetraploid and hexaploid species (Fig. 2, bands c–h). The nomenclature for bands used here differs from that used in previously published studies (Konarev et al., 2000a) because of improved resolution. Wild forms of *H. annuus* (Fig. 2, tracks 6–9) contained the majority of bands present in cultivated *H. annuus* while accessions of wild diploid perennial species (e.g. Fig. 2, tracks 10–12) differed from the annual species in the absence of bands c to f. The majority of tetraploid species also lacked these bands with only one *H. hirsutus* acc. having band d (Fig. 2, track 14) and *H. strumosus* having a band at position g not present in *H. annuus* (Fig. 2, track 17). In hexaploid species, bands c and f were present in accession of *H. resinosus* (Fig. 1, track 22) and bands e and f in *H. multiflorus* (Fig. 1, track 23) while *H.*

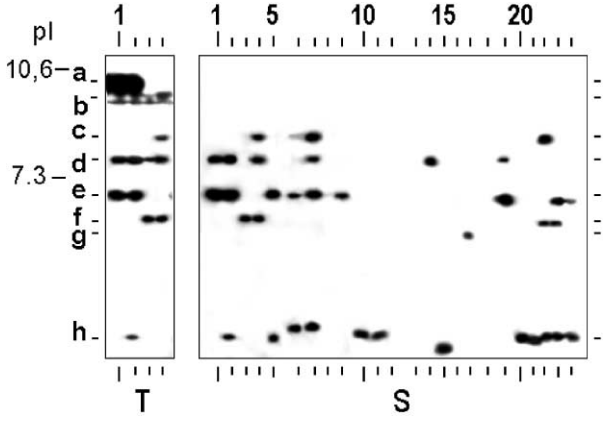

Fig. 2. Polymorphism of proteinase inhibitors in seeds of cultivated and wild *Helianthus* species. Water-soluble proteins were separated by IEF in the pH range 3–10. Gelatin replicas were developed using trypsin (T) and subtilisin (S). a–h, main positions of inhibitor bands. Slots **1–9** and **19**, *H. annuus*: **1** and **19**, line VIR-130; **2**, VIR-104; **3**, VIR-369; **4**, VIR-648b; **5**, var. Sunbred 246; **6–9**, wild forms: **6**, ssp. *annuus*; **7–9**, ssp. *lenticularis*, various acc. **10–12**, diploid perennial species: **10**, *H. giganteus*; **11**, *H. salicifolius*; **12**, *H. occidentalis* ssp. *occidentalis*. **13–18**, tetraploid species: **13** and **14**, *H. hirsutus*, acc.; **15** and **16**, H. *decapetalus*, acc.; **17**, *H. strumosus*; **18**, *H. laetiflorus*. **20–24**, hexaploid species: **20**, *H. californicus*: **21**, *H. rigidus*; **22**, *H. resinosus*; **23**, *H. multiflorus*; **24**, *H. tuberosus*.

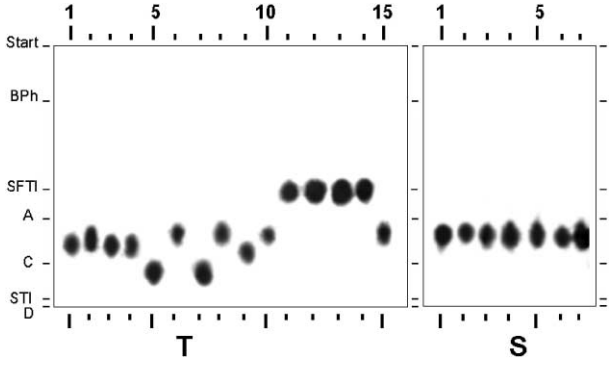

Fig. 3. Variation in the masses of native proteinase inhibitors in seeds of Compositae species revealed by thin layer gel filtration. Water-soluble seed proteins were loaded on gels in volumes of 0.5–2 μl and separated by thin layer gel filtration in Sephadex G-50 (Superfine). Gelatin replicas were developed using trypsin (T) and subtilisin (S). Positions of markers; BPh, bromphenol blue (M_r 692); SFTI, main sunflower TI (M_r 1514); A, aprotinin (TI from bovine lung, M_r 6500); C, cytochrome c (M_r 12,400); STI, soybean TI (M_r 21,000); D, Dextran Blue. **T.1**, *Silphium perfoliatum*; **2**, *Cosmos bipinnatus*; **3**, *Zinnia elegans*; **4**, *Rudbeckia laciniata*; **5**, *Gaillardia aristata*; **6**, *Bidens pilosa*; **7**, *Eclipta prostrata*; **8**, *Senecio viscosus*; **9**, *Solidago virgaurea*; **10**, *Centaurea cyanus*; **11**, *Helianthus. annuus* VIR-104; **12**, H. *tuberosus*; **13**, *Tithonia diversifolia*; **14** and **15**, *H. annuus* inhibitors purified by affinity chromatography and HPLC: **14**, SFTI; **15**, fraction T/SI. **S.1**, *H. annuus*, fraction T/SI; **2**, *Carthamus tinctorius*; **3**, *Centaurea cyanus*; **4**, *C. triumfettii*; **5**, *Lactuca sativa*; **6**, *Taraxacum officinale*; **7**, *Senecio viscosus*.

tuberosus had a weak band e (Fig. 1, track 39). All studied accessions of hexaploid species had band h.

TLGF revealed that the major forms of T/SI found in seeds of representatives of the three subfamilies had similar masses (Fig. 3). The native T/SI from *Centaurea*, *Carthamus*, *Lactuca*, *Taraxacum*, *Cosmos* and *Bidens* species and purified *H. annuus* T/SI (a fraction containing a mixture of IEF bands d and e in Fig. 2; Fig. 3, T, track 15 and S, track 1) had mobilities on TLGF between aprotinin (M_r 6,500) and cytochrome c (M_r 14,400) (the activity of the native sunflower T/SI was not sufficiently high for detection after TLGF). TI from different sources had M_r ranging from the lowest (SFTI-1) in *H. annuus*, *H. tuberosus* and *T. diversifolia* (Fig. 3, T, tracks 11–13), which migrated between bromophenol blue (M_r 692) and aprotinin (M_r 6,500) to the highest in *Gailardia aristata* (Fig. 3, track 5) and *Eclipta prostrata* (Fig. 3, track 7) with TI from *Silphium perfoliatum*, *Zinnia elegans*, *Rudbeckia laciniata* and *Solidago virgaurea* (Fig. 3, tracks 1, 3, 4 and 9, respectively) being intermediate in mass.

2.2. Purification and characterization of inhibitors from seeds

2.2.1. Safflower (*C. tinctorius var. Goldtuft*)

Proteinase inhibitors were purified using affinity chromatography on trypsin–Sepharose followed by RP–HPLC (Fig. 4A). Analytical IEF combined with visualisation using the gelatin replicas method showed that the major HPLC peak contained T/C/SI and the smaller peak SI while weak TI were present in a minor fraction eluted before the other inhibitors. Micropreparative IEF followed by a second RP–HPLC separation gave fractions containing T/C/SI and SI (Fig. 4B).

An initial attempt to sequence the major safflower inhibitor (T/C/SI) failed because the *N*-terminal amino acid residue was blocked. The inhibitor was therefore cleaved with subtilisin and cysteine residues blocked by alkylation with *N*-isopropyl iodacetamide (NIPIA). HPLC revealed five peaks in the reaction mixture, not shown, with peak B corresponding to subtilisin and peak C to the intact inhibitor (M_r 7555, Table 2) modified with one molecule of NIPIA (M_r 199). The masses of peaks D (M_r 4848) and E (M_r 2923) were consistent with their production by cleavage of the modified inhibitor at the active site. The N-terminal sequence of peptide E was determined for 10 residues (Table 3) and comparison with protein sequence databases revealed homology with sequences adjacent to the P_1 residues at the active sites of members of the potato I chymotrypsin inhibitor family. The safflower SI gave an identical *N*-

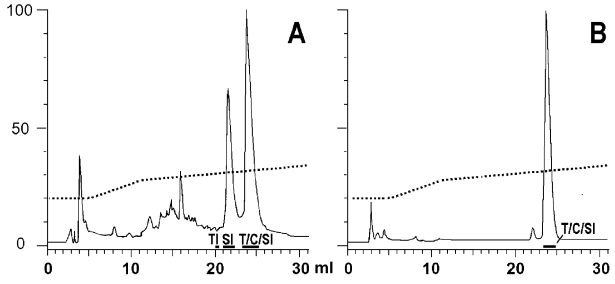

Fig. 4. Purification of proteinase inhibitors from safflower (*Carthamus tinctorius* seeds). **A.** Proteins eluted from a chymotrypsin–Sepharose column with 0.015 M HCl were freeze dried and separated by RP–HPLC on a Vidac C18 column with an acetonitrile gradient. TI, SI and T/C/SI, inhibitor types detected in corresponding fractions with the gelatin replica method. The major protein fraction containing T/C/SI was separated by micropreparative IEF in Servalyt Precotes pH 3–10 gels and the major band eluted and subjected to a second RP–HPLC separation (B).

Table 2
Molecular masses of proteinase inhibitors purified from seeds of some species of the Compositae

Species	Type	M_r	Species	Type	M_r
Carthamus tinctorius	T/C/SI	7555	*Gaillardia aristata*	T/SI	7449
	Peptide C	4848		TI (7 IEF	10238
	Peptide D	7754		bands)	11372
	Peptide E	2923		TI (IEF band)	14753
	SI	7572	*Silphium perfoliatum*	TI	11439
Centaurea cyanus	T/C/SI	7606	*Helianthus annuus*	TI (h14)	1514
Taraxacum officinale	T/SI (a)	7482		TI (h4)	1531
	T/SI (b)	7542		TI (h9)	1531
Zinnia elegans	TI (z1)	11350		T/SI (h25)	7723
Cosmos bipinnatus **T**	T/C/SI (a)	7680		TI (h29)	7613
T	T/C/SI (d)	7722		T/SI (h30)	7618
T	T/SI (e)	7740		T/SI (h31)	7596
C	C/SI (f)	7775		T/SI (h33)	7610
C	C/SI (g)	7792		T/SI (h35)	7593
C	C/SI (h)	7670			

a, b, h25, z, designation of IEF bands; **T** and **C**, affinity ligands used to isolate inhibitors (trypsin and chymotrypsin).

Table 3
Alignment of the sequences linked to the P_1 residue of proteinase inhibitors from safflower (*Carthamus tinctorius*) with those of two inhibitors from *Cosmos bipinnatus* and some other representatives of the potato inhibitor I family

Organism	Inhibitor	Res. Nos.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Fagopyrum esculentum[1]	PI	46	D	L	R	C	D	R	V	W	V	F
Cosmos bipinnatus[2]	T/SI (e)	?	D	L	R	C	D	R	V	W	V	F
Amaranthus hypochondriacus[3]	TI	46	D	F	R	C	D	R	V	W	V	V
Carthamus tinctorius[4]	Peptide E	1	D	F	R	C	D	R	V	W	V	W
Carthamus tinctorius[5]	SI	?	D	F	R	C	D	R	V	W	V	W
Momordica charantia[6]	TI	45	D	F	R	C	D	R	V	R	V	W
Arabidopsis thaliana[7]	PI	50	D	F	R	C	D	R	V	R	V	F
Cosmos bipinnatus[2]	C/SI (g)	?	D	Y	R	C	D	R	V	R	V	F
Lycopersicon peruvianum[8]	PI	88	D	F	R	C	D	R	V	R	L	F
Nicotiana sp.[9]	PI	71	D	L	R	C	D	R	V	R	L	F
Hordeum vulgare[10]	CI 2	61	E	Y	R	I	D	R	V	R	L	F
Solanum tuberosum[11]	CI A	47	D	F	R	C	N	R	V	R	L	F
Solanum tuberosum[12]	CI C	62	D	Y	R	C	N	R	V	R	L	F
Hirudo medicinalis (leech)[13]	Eglin C	46	D	L	R	Y	N	R	V	R	L	F
Cucurbita maxima[14]	PFTI	44	D	Y	R	P	N	R	V	R	L	F
Hordeum vulgare[15]	CI 1	54	N	F	N	P	N	R	V	F	L	L

Residues identical to those of *C. tinctorius* SI are shown in black boxes; conserved substitutions are in grey boxes: nonrelated residues are in black letters in white boxes. Res. nos., numbers of first residues of corresponding sequences; *1–10*, numbering of residues in *Carthamus tinctorius* peptide E. [1]Belozersky et al. (1995); [2]this paper (Fig. 5); [3]Valdes-Rodriguez et al. (1993); [4]this paper; [5]this paper (Fig. 4A); [6]Miura and Funatsu (1995); [7]Lin et al. (1999); [8]Wingate et al. (1989); [9]Fujita et al. (1993); [10]Peterson et al. (1991); [11]Richardson (1974); [12]Richardson and Cossins (1974); [13]Seemuller et al. (1980); [14]Murray and Christeller (1995); [15]Svendsen et al. (1982).

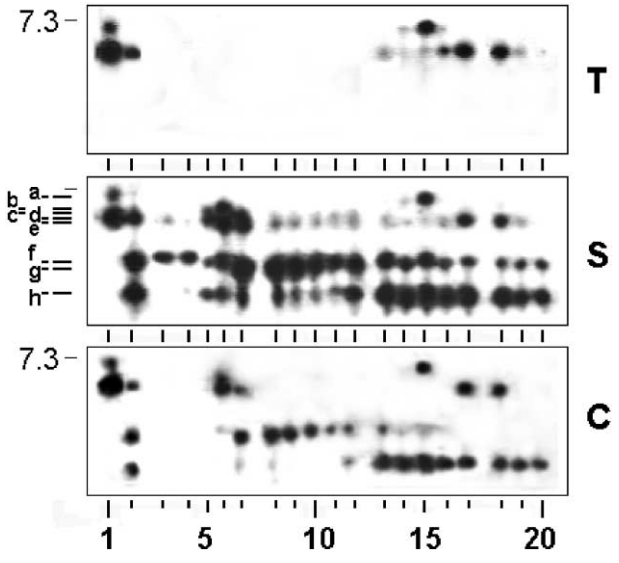

Fig. 5. Fractionation of serine proteinase inhibitors from *Cosmos bipinnatus* seeds using affinity chromatography, RP–HPLC and analytical IEF. Proteins eluted from chymotrypsin–Sepharose with 0.015 M HCl were freeze dried and separated by RP–HPLC with an acetonitrile gradient. Fractions collected each 0.5 min were analyzed by IEF. Inhibitors of trypsin (T), subtilisin (S) and chymotrypsin (C) were detected using three gelatin replicas obtained from the same gel. 1 and 2, seed proteins eluted from trypsin– and chymotrypsin–Sepharose, respectively. 3–20, Fractions collected from RP–HPLC: tracks 3 and 4 are 0.5 ml fractions from 27.5 and 28 ml; tracks 5–20 are 0.5 ml fractions from 30.5 to 38 ml. a–h, Positions of some inhibitor components.

terminal sequence to peptide E. The mass of the SI (7572) differed from that of the intact T/C/SI (7555) by only 17 units which is consistent with its origin by cleavage of the T/C/SI at the active site during purification.

2.2.2. Cosmos (*Cosmos bipinnatus*)

Fractions enriched for inhibitors by affinity chromatography on trypsin–Sepharose and chymotrypsin–Sepharose columns were separated by IEF (Fig. 5, tracks 1 and 2) and inhibitors of trypsin (T), subtilisin (S) and chymotrypsin (C) detected using gelatin replicas. All the inhibitors showed activity against subtilisin and trypsin and/or chymotrypsin. The fractions from affinity chromatography were separated by RP–HPLC (see Fig. 6, tracks 3–20 for the chymotrypsin column, fractions from the trypsin column are not shown) and selected fractions from the chymotrypsin column (Fig. 5, tracks 6, 9, 14) were separated by micropreparative IEF followed by a second separation by RP–HPLC. This resulted in the purification of six inhibitors corresponding to bands a and d–h in Fig. 5. They had similar M_r ranging from 7670 to 7792, the differences being consistent with bands e (T/SI) and g (C/SI) being derived from proteolytic cleavage of bands d (T/SI) and f (C/SI), respectively (the differences being 18 and 17 mass units). Sequencing of components e (T/SI), g (C/SI) and a (T/C/SI) showed homology to members of

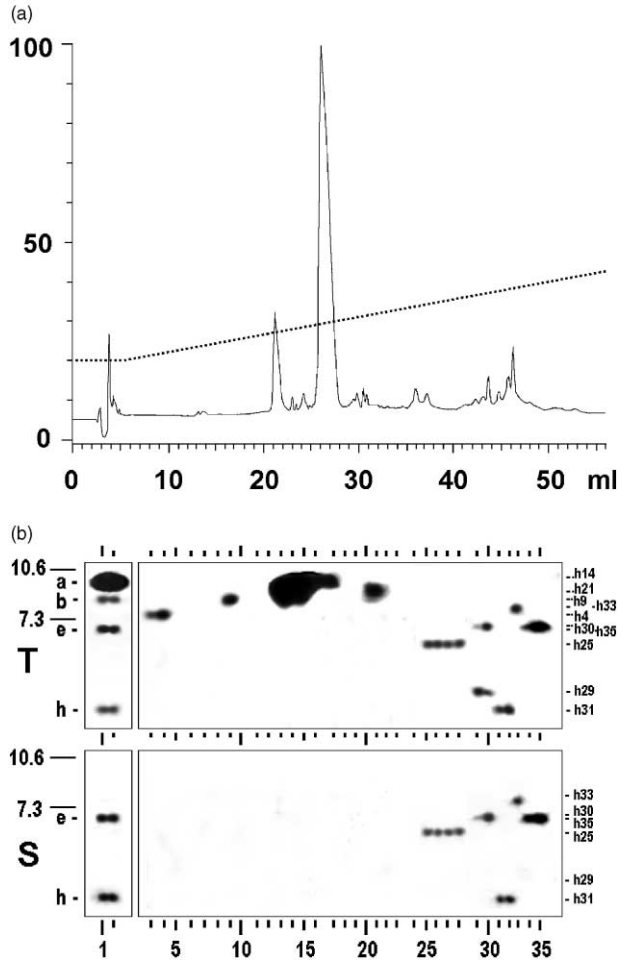

Fig. 6. Purification of proteinase inhibitors from sunflower (*Helianthus annuus*, var. Alsan) seeds. The fraction enriched for inhibitors using affinity chromatography on trypsin–Sepharose was subjected to RP–HPLC (a) with a gradient of acetonitrile (·····). Fractions were analyzed by IEF followed by detection of inhibitors using the gelatin replicas method (b). Replicas from the same gel were developed with trypsin (T) and subtilisin (S). a, b, e and h, Positions of native inhibitors in sunflower accessions (Fig. 2); h4-h35, positions of purified inhibitors numbered according to fractions eluted. 1 and 2, native inhibitors, extracted with water; 3–35, HPLC-fractions tracks 3–23 are 0.5 ml fractions from 21.5–31.5 ml; tracks 24–28 are 0.5 ml fractions from 36–38 ml; tracks 29–35 are 0.5 ml fractions from 44–47 ml.

the potato inhibitor I family, with e and g giving sequences related to the active site region (cf safflower peptide e) (Table 3) while component a gave a sequence related to the *N*-termini (Table 4). This indicates that the sequences determined for components e and f resulted from cleavage at the active site during purification, the *N*-termini of the intact proteins being blocked.

2.2.3. Sunflower (*H. annuus*)

Affinity chromatography on trypsin–Sepharose followed by RP–HPLC, micropreparative IEF and a second RP–HPLC separation was used to purify several TI and T/SI components (Fig. 6), the former being represented by components h4, h9, h14 and h21. Differences in M_r (Table 2) indicated that h4 and h9 were both proteolytically cleaved forms of the main TI component h14, corresponding to SFTI-1, and this was confirmed for h4 which showed sequence corresponding to the active site of SFTI-1 (Table 5).

The other sunflower inhibitors had M_r ranging from 7596 to 7723 (Table 2), with one TI and five T/SI. Comparison of their M_r indicates that h29 (TI) could be derived from h31 (T/SI) by proteolysis leading to loss of activity against subtilisin. Similarly h33 (T/SI) could be derived from h35 (T/SI). No sequences of these components were determined but h35 (T/SI) was found to be *N*-terminally blocked.

2.2.4. Other species

The major TI in seeds of *Zinnia elegans* had a mass of 11,350. *N*-terminal sequencing for 20 residues showed no relationship to other inhibitors of the Compositae but limited homology to Bowman–Birk type inhibitors from various sources (Table 6). This sequence was therefore, confirmed using a second sample prepared using affinity chromatography, RP–HPLC and micropreparative SDS electrophoresis followed by transfer of proteins to PVDF membrane.

Similar analyses showed one T/SI component (M_r 7449) and at least 14 TI components in seeds of *Gaillardia aristata*. Seven of these were purified showing that most had M_r ranging from 10,238 to 11,372 with one of M_r 14,753. *N*-terminal sequencing did not give clear results, possibly due to microheterogeneity of the fractions.

Analyses of *Silphium perfoliatum* seeds showed an *N*-terminally blocked TI of M_r 11,439 while a T/C/SI of M_r 7606 was demonstrated in *Centaurea cyanus* and two possibly related T/SI of M_r 7482 and 7542 in *Taraxacum officinale* (dandelion).

3. Discussion

Results from mass determination (Table 2) and amino acid sequencing (Tables 3–6) indicate that the proteinase (C, S, T) inhibitors present in seeds of the Compositae fall into three major classes, as summarised in Table 7.

The most widespread group appears to be related to the potato inhibitor I family, with M_r ranging from 7450 to 7800 and activity against two or three serine proteinases (T/SI, C/SI, T/C/SI). These inhibitors can be expected to be active against extracellular proteinases of pathogenic fungi such as *Sclerotinia sclerotiorum* and *Botrytis cinerea* and may, therefore, be involved in defense mechanisms. In terms of evolutionary relationships, this family of inhibitors occurs in only low amounts with one or two isoforms being present in each species in the Carduiodeae, which is considered to be one of the most primitive subfamilies of the Compositae (Bremer, 1996), but is highly heterogeneous in some

Table 4
Alignment of the *N*-terminal amino acid sequence of the T/C/SIa proteinase inhibitor from seeds of *Cosmos bipinnatus* with those of some members of the potato inhibitor I family

Organism	Inhibitor	Residue Nos.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*		
Cosmos bipinnatus[1]	T/C/SI (a)	1	A	R	E	C	K	G	K	R	A	W		
Amaranthus hypochondriacus[2]	TI	1	A	R	E	C	P	G	K	Q	E	W	P	E
Linum usitatissinum[3]	TI	2	S	R	R	C	P	G	K	N	A	W	P	E
Fagopyrum esculentum[4]	BWI-1	1	L	R	Q	C	S	G	K	Q	E	W	P	E
Momordica charantia[5]	TI	1	-	S	R	C	Q	G	K	S	S	W	P	Q
Solanum tuberosum[6]	CI-I, C	2	E	F	E	C	K	G	K	L	Q	W	P	E
Hordeum vulgare[7]	CI-1	19	I	G	A	S	G	A	K	F	S	W	P	E
Hordeum vulgare[8]	CI-2	16	G	D	R	Q	N	Q	K	T	E	W	P	E
Hirudo medicinalis[9]	Eglin C	1	T	E	F	G	S	E	L	K	S	F	P	E

Residues identical to those of *Cosmos bipinnatus* are in black boxes and conservative changes are in grey boxes. Residue Nos., numbers of first residues of corresponding sequences; *1–10*, numbering of *C. bipinnatus* inhibitor residues. [1] This paper (Fig. 5); [2]Valdes-Rodriguez et al. (1993); [3]Cierpicki et al. (1999); [4]Belozersky et al. (1995); [5]Miura and Funatsu (1995); [6]Richardson and Cossins (1974); [7]Williamson et al. (1988); [8]Peterson et al. (1991); [9]Seemuller et al. (1980).

Table 5
Comparison of the amino acid sequences of sunflower (*Helianthus annuus*) TI

Vigna radiata[1]	TI	22	**C**RCTK**SIPPQC**HCA	35
H. annuus[2]	SFTI-1		<**GRCTK**SIPPICFPD>	
H. annuus[3]	TI (h4)		**SIPPICFPDGRCTK**	

"<...>" Cyclic bond. Bold letters indicate residues identical to those of SFTI-1. Underlined letters indicate identical segments of SFTI-1 and h4. [1]Wilson and Chen (1983); [2]Luckett et al. (1999); [3]this paper (Fig. 6).

Table 6
Alignment of the amino acid sequence of the ZTI proteinase inhibitor from seeds of *Zinnia elegans* with those of some members of the Bowman-Birk proteinase inhibitor family

	Residue Nos.	*1*				*5*				*10*						*15*					*20*			
Zinnia ZTI[1]	1		P	W	E	Q	C	R	S	Q	H	A	-	-	I	E	K	L	N	H	X	Q	M	H		
Vigna radiata BBI[2]	1			E	P	C	C	D	S	-	C	R	-	-	C	T	K	S	I	P	P	Q	C	H	C	A
Consensus legume BBI[3]							C	R	S	-		A	-	-	V	D	K		N			Q		H		
Hordeum vulgare D1[4]	5	R	P	W	K	C	C	D	E	-	A	V	-	-	C	T	R	S	I	P	P	I	C	T	C	M
Coix lachryma-jobi[5]	5	R	F	W	E	C	C	D	I	-	A	M	-	-	C	T	R	S	I	P	P	I	C	R	C	V
Oryza sativa D1[6]	62	K	P	W	K	C	C	D	N	-	H	E	R	L	P	T	K	T	N	P	P	Q	W	R	C	N
Consensus cereal BBI[7]			P	W	E		C		N	-	H		-	-			K	M	N			Q				
Consensus BBI			P	W	E		C	R	S	-	H	A	-	-	V	D	K	M	N			Q		H		
Some other conservative or different BBI residues[8]					D			K	A			S					R		D	R	A		Y	F		

Residues identical to those of ZTI are in black boxes while conservative changes are in grey boxes. Dashes indicate spaces introduced to maximise similarity. *1*, *5*, *10*, *15*, *20* are numbers of residues in ZTI; 1, 1, 5, 5, 62 are first residues of the sequences shown. [1]This paper, [2]Li et al. (1994); [3,7,8]Internet source http://www.ncbi.nlm.nih.gov/BLAST; [4]Nagasue et al. (1988); [5]Ary et al. (1988); [6]Chen and Chen (1997). The *Hordeum vulgare* and *Oryza sativa* sequences are from the *N*-terminal domains of duplicated BBI.

groups of the advanced subfamily Asteroideae, especially in the tribe Heliantheae (*Helianthus*, *Cosmos*, *Bidents*, *Dahlia*). In sunflower (*Helianthus annuus*) these inhibitors can be used to differentiate lines and varieties while the absence of some components (c–f) from perennial diploid species with the A genome confirms the significant evolutionary separation of the A and B genomes present in diploid species. It is interesting that diploid perennial species of *Helianthus* also lack several helianthinin (11S storage globulin) components characteristic of annual forms (Anisimova, 1996).

Partial amino acid sequences were determined for the *N*-terminus of one inhibitor related to the potato inhibitor I family (*Cosmos* T/C/SIs) but the T/C/SI from *Carthamus* was *N*-terminally blocked. However, sequences corresponding to the active site residues (starting at residue P'_1) were determined for a second component from *Carthamus* (SI) and two components from *Cosmos*. In each case these appeared to be derived from cleavage of an N-terminally blocked inhibitor at the active site during affinity purification, as summarised in Table 7. However, in some cases the putative cleaved and native

Table 7
Probable identity and relationships of proteinase inhibitors purified from seeds of some species of the Compositae

		Intact forms			Proteolytically cleaved forms		
Inhibitor family	Species	Designation/specificity	M_r	Sequence determined	Designation/specificity	M_r	Sequence determined
Potato inhibitor I	*Carthamus*	T/C/SI	7555	blocked	SI	7572	Active site
	Cosmos	T/C/SI a	7680	N-terminus			
		T/SI d	7722	nd	T/SI e	7740	Active site
		C/SI f	7775	nd	C/SI g	7792	Active site
	Centaurea	T/C/SI	7606	nd			
	Taraxacum	T/SI	7482	nd			
		T/SI	7542	nd			
	Gaillardia	T/SI	7449	nd			
	Cosmos	C/SI	7670	nd			
	Helianthus	T/SI h31	7596	nd	TI h29	7613	nd
		T/SI h35	7593	nd	T/SI h33	7610	nd
		T/SI h25	7723	nd			
Bowman–Birk	*Helianthus*	TI h14	1514	blocked	TI h41 (SFTI-1)	1531	Active site
Novel/unidentified	*Zinnia*	ZTI	11,350	N-terminus			
	Gaillardia	TI	10,238	nd			
			11,372	nd			
	Silphium	TI	11,439	nd			
	Gaillardia	TI	14,753	nd			

nd, Not determined.

The proteolytically cleaved forms are proposed to be derived from the corresponding intact forms by proteolysis at the active site during purification by affinity chromatography.

inhibitors showed different specificities (the SI and T/C/SI of *Carthamus* and the TIh29 and T/SIh31 of *Helianthus*) which indicates that cleavage may result in subtle differences in the active site geometry. The sequences determined are clearly homologous with those of other members of the potato inhibitor I family with particularly high identify of TI (e) from *Cosmos* with a PI from *Fagopyrum esculentum*, T/C/SI from *Carthamus* and T/C/SI (a) from *Cosmos* with TI from *Amaranthus hypochondriacus* and C/SI (g) from *Cosmos* with PI from *Arabidopsis thaliana*. This may indicate the existence of several major variants of inhibitors of the potato inhibitor I family (at least in parts of sequences in regions near reactive sites) which are present in plants and animals. Analyses of the *Carthamus* T/C/SI also indicated the presence of only one cysteine residue which is consistent with other members of the potato inhibitor I family. The greater heterogeneity of TSI in the Helianthеae in comparison with the Cardueae may be due to an amphidiploid origin of their apparently diploid genomes (Sossey-Alaoui et al., 1996). A high level of polymorphism and variability in T/C/SI belonging to the potato I inhibitor family is also characteristic of other plant groups, for example in wheat (*Triticum aestivum*) (Poaceae) where such inhibitors can be used for varietal identification (Konarev, 1988).

Two other types of trypsin inhibitors found in seeds of the Compositae may be related to the Bowman–Birk inhibitor (BBI) family. The first type comprises a unique cyclic peptide inhibitor of M_r 1514 (SFTI-1) previously characterized from seeds of sunflower (Luckett et al., 1999). The present study resulted in the isolation of two peptides (h4 and h9) with identical masses but different *p*I and hydrophobicities which appeared to be derived from SFTI-1 by proteolysis. Sequence analysis of h4 confirmed that it was derived from SFT1 by cleavage between the P_1 and P'_1 residues (Table 5) and a related peptide appeared to also be present in seeds of *Tithania diversifolia* but not in other species. These inhibitors appear to be derived from the trypsin-reactive loop of a BBI, perhaps by the action of a transpeptidase enzyme to introduce a cyclic structure (Luckett et al., 1999). Cyclic trypsin inhibitors with M_r higher than 3,000 belonging to a different protein family were recently reported in some Cucurbitaceae (Felizmenio-Quimio et al., 2001) and it is possible that inhibitors with cyclic structures are widely distributed in plants. Other TI detected in seeds of sunflower in this (e.g. band

b) and previous studies (TI of M_r 2,500, Konarev et al., 2000a) have not been characterized in sufficient detail to determine their relationships.

A novel type of TI with M_r of 11,350 (ZTI) was identified in seeds of *Zinnia*, while inhibitors of M_r 10,238 and 11,372 in *Gaillardia* and of 11,439 in *Silphium* could belong to the same family. Sequence comparisons indicated that these could be related to the Bowman-Birk inhibitors (BBI) which are widely distributed in seeds of legumes and other species including cereals. Most BBI have M_r of 8000–9000 but duplicated forms of M_r about 14,000 are present in some cereal seeds (Shewry, 1999). The *N*-terminal sequence of ZTI contained a pentapeptide motif (PWEQC) which was similar to sequences close to the *N*-termini of BBI from cereals such as barley (PWECC), rice (PWEDC) and Job's tears (*Coix lachryma-jobi*) (PWECC) (see Table 6). However, this motif is not present in Bowman-Birk inhibitors from legumes or other dicotyledonous species and similar motifs also occur in pectolytic enzymes and cereal prolamins (Okita et al., 1985; Robert et al., 1993). The ZTI also lacked cysteines in positions equivalent to those that stabilize the first reactive loop (i.e. Cys9 and Cys17 in the *Vigna radiata* sequence shown in Table 6). However, these cysteines are also absent from the first domain of the duplicated D1 inhibitor of rice (Table 5). Furthermore, the ZTI has a basic residue (Lys) in the position corresponding to the active site (Lys11 in *V. radiata*, Lys or Arg in other BBI). It is possible, therefore, that the ZTI is distantly related to the BBI but further studies are required to confirm this. A further TI (M_r 14,753) was present in *Gaillardia* seeds but was not characterized in detail.

The studies reported here confirm that seeds of the Compositae are rich sources of proteinase inhibitors, belonging to at least three separate families. Although their biological roles remain unknown it can be speculated that at least some are involved in conferring resistance to pests and pathogens. It will, therefore, be of interest to determine their activities against fungal pathogens and extracellular proteinases in relation to their potential exploitation to confer resistance in transgenic plants.

4. Experimental

4.1. Seeds and leaves

Seeds of various species of the *Compositae* were obtained from the world collection of the Vavilov Institute of Plant Industry (VIR, St. Petersburg, Russia), the herbarium of the Komarov Botanical Institute (St. Petersburg), Long Ashton Research Station (Bristol, UK), "Herbeseed" (UK) and "Nickerson-Zwaan" (UK) or were collected by the authors.

The Compositae was represented by 128 species from 65 genera of the subfamilies Carduoideae (9 genera, 26 species), Cichorioideae (10 genera, 17 species), and Asteroideae (46 genera, 85 species). The main taxa were specified according to Bremer (1996) with many species being represented by several accessions (acc.) of different origin.

4.1.1. Subfamily Carduoideae

Safflower *Carthamus* L. species with **2n = 24**: *C. tinctorius* L. (25 acc. from Afghanistan, Czechoslovakia, Egypt, Ethiopia, Germany, Hungary, Mexico, Russia, Tadzhikistan and Uzbekistan with VIR catalog Nos. k-4, k-8, k-11a, k-11, k-14, k-18, k-21, k-26, k-62, k-64, k-66, k-69, k-72, k-92, k-95, k-120, k-266, k-404, k-405, k-410, k-442, k-503, k-506, k-569 and var. Goldtuft from UK); *C. oxyacanthus* Bieb (India); *C. palaestinus* Eig ex K.H.Rechinger (Israel). **2n = 20:** *C. glaucus* Bieb (Israel and USA). **2n = 44:** *C. lanatus* L. (Germany and Uzbekistan). *Arctium* L.: *A. tomenntosum* Mill., *A. lappa* L. *Carduus* L.: *C. acanthoides* L., *C. crispus* L., *Centaurea*: L.: *C. cyanus* L., *C. jacea* L., *C. scabiosa* L., *C. triumfettii* All. *Cirsium* Mill.: *C. arvense* (L.) Scop., *C. vulgare* (Savi) Ten. *Cousinia* Cass.: *C. badghysi* Kult., *C. microcarpa* Boiss., *C. olgae* Regel & Schmalh., *C. raddeana* C. Winkl., *C. schistoptera* Juz. *Cynara* L.: *C. scolymus* L. *Saussurea* DC.: *S. amara* (L.) DC., *S. parviflora* (Poir.) DC., *S. pulchella* (Fish.) Fisch., *S. salicifolia* (L.)DC. *Serratula* L.: *S. coronata* L.

4.1.2. Subfamily Cichorioideae, tribe Lactuceae Cass.

Cichorium L.: *C. intybus* L. *Hieracium* L.: *H. aurantiacum* L. *hypochaeris* L.: *H. radicata* L. *Lactuca* L.: *L. sativa* L. [23 varieties (var.) from Canada, China, Germany, Hungary, Italy, Japan, Kazakhstan, The Netherlands, Russia, Spain and USA], *L. serriola* L., *L. livida* Boiss. et Reut., *L. quercina* L., *L. tatarica* (L.) C.A. Mey. *Lapsana* L.: *L. communis* L. *leontodon* L.: *L. autumnalis* L. *Scorzonera* L.: *S. hispanica* L. *Sonchus* L.: *S. arvensis* L., *S. oleraceus* L. *Taraxacum* Wigg.: *T. officinale* Wigg. (acc. from St. Petersburg region, Germany and "Herbeseed"), *T. kok-saghyz* Rodin (k-1129), *T. hybernum* Stev. (k-1 and k-4, Russia and k-7, Germany): *Tragopogon* L.: *T. pratensis* L.

4.1.3. Subfamily Asteroideae, tribe Astereae Cass.

Aster L.: *A. tripolium* L. *Bellis* L.: *B. perennis* L. *Galatella* Cass.: *G. dracunculoides* (Lam.) Nees (leaf). *Grindelia* Willd.: *G. integrifolia* DC. *Erigeron* L.: *E. canadensis* L., *E. caucasicus* Stev., *E. uniflorus* L. *Solidago* L.: *S. virgaurea* L.

4.1.3.1. Tribe Anthemideae Cass. Achillea L.: *A. ptarmica* L., *A. millefolium* L. *Artemisia* L.: *A. vulgaris* L. *Chrysanthemum* L.: *C. coronarium* L., *C. carinatum* Schousb. *Leucanthemum* Mill.: *L. maximum* (Ram.)

DC. *Matricaria* L.: *M. matricarioides* (Less.) Porter. *Pyrethrum* Zinn: *P. partenium* (L.) Smith, *P. balsamita* (L.) Willd. *Tanacetum* L.: *T. vulgare* L. *Tripleurospermum* Sch.-Bip.: *T. inodorum* (L.) Sch.-Bip.

4.1.3.2. Tribe Senecioneae Cass. Doronicum L.: *D. pardalianches* L. *Emilia* Cass.: *E. sonchifolia* DC.; *Erechtites* Rafin.: E. *valerianifolia* (Wolf) DC.; *Ligularia* Cass.: *L. dentate* (A. Gray) Hara. *Senecio* L.: *S. congestus* (R. Br.) DC., *S. cineraria* DC., *S. thyrsophorus* C. Koch, *S. sakalavorum* Humbert, *S. viscosus* L.

4.1.3.3. Tribe Eupatorieae Cass. Ageratum L.: *A. houstonianum* Mill., *A. conyzoides* L. *Eupatorium* L.: *E. cannabinum* L., *E. odoratum* L.

4.1.3.4. Tribe Tageteae Cass. Tagetes L.: *T. erecta* L., *T. patula* L.

4.1.3.5. Tribe Calendulae Cass. Calendula. L.: *C. officinalis* L. *Dimorphotheca* Moench: *D. pluvialis* (L.) Moench.

4.1.3.6. Tribe Inuleae Cass. Helichrysum Mill.: *H. armenium* D.C., *H. bracteatum* (Vent) Andr.; *Inula* L.: *I. helenium* L.

4.1.3.7. Tribe Heliantheae Cass. Alvordia Brandegee: *A. fruticosa* Brandegee. *Arnica* L.: *A. alpina* (L.) Olin, *A. iljinii* (Maguire) Iljin. *Bidens* L.: *B. tripartita* L., *B. pilosa* L., *B. radiata* Thuill. *Coreopsis* L.: *C. tinctoria* Nutt. *Cosmos* Cav.: *C. bipinnatus* Cav. (acc. from St. Petersburg region, ''Nickerson - Zwaan'' and wild form from Kenya), *C. caudatus* Kunth: *Dahlia* Cav.: *D. pinnata* Cav. *Echinacea* Moench: *E. purpurea* (L.) Moench. *Eclipta* E. Mey.: *E. prostrata* L. *Flourensia* DC.: *F. cernua* DC. *Galinsoga* Riuz et Pav.: *G. parviflora* Cav. *Gaillardia* Foug.: *G. aristata* Pursh. *Rudbeckia* L.: *R. hirta* L., *R. laciniata* L., *R. speciosa* Wend. *Silphium* L.: *S. perfoliatum* L. *Zinnia* L.: *Z. elegans* Jacq., *Z. angustifolia* Kunth.

4.1.3.8. Subtribe Helianthinae Dumort. Enceliopsis A. Nels.: *E. argophylla* (D.C. Eaton) A.Nels. *Simsia* Pers.: *S. calva* A.Gray, *S. annectens* S.F.Blake, *S. eurylepis* S.F.Blake. *Tithonia* Desf. ex Juss.: *T. diversifolia* (Hemsl.) A.Gray (3 acc. from Madagascar, Seashell islands, and ''Nickerson-Zwaan''); *Wedelia* Jacq.: *W. spilanthoides* F. Muell. *Viguiera* Kunth: *V. discoidea* (Griseb.) Blake. *Helianthus* L.: 10 var. and 70 lines of cultivated *H. annuus* L.; Wild diploid annual forms *H. annuus* ssp. *annuus* and *H. annuus* ssp. *lenticularis* (Dugl. ex Lindl.) Cockerell. Diploid perennial species: *H. giganteus* L., *H. salicifolius* A. Dietr., *H. occidentalis* Riddell ssp. *occidentalis, H. occidentalis* var. *plantagineus* Torr. et Gray, *H. mollis* Lam., *H. microcephalus* Torr.et A.Gray, *H. maximilianii* Schrad., *H. grosseserratus* M.Martens, *H. divaricatus* L. Tetraploid species: *H. hirsutus* Rafin., *H. decapetalus* L., *H. strumosus* L., *H. laetiflorus* Pers. Hexaploid species: *H. californicus* DC., *H. multiflorus* L., *H. resinosus* Small, *H. rigidus* (Cass.) Desf., *H. tuberosus* L.

Seed and leaf material from herbarium specimens collected after 1960 and fresh leaves stored at −20 °C were used in addition to seeds.

4.2. Proteins

4.2.1. Extraction

For analytical separation, proteins were extracted from ground seeds or leaves by shaking with water (1:4 w/v) for 1 h at 20 °C. For preparative separation, milled seeds were defatted with hexane and extracted with water (1:10 w/v). The mixture was centrifuged for 30 min at 20,000×*g* at 20 °C.

4.2.2. Analytical separation

Protein fractions were separated by isoelectric focusing (IEF) in Servalyt Precotes pH 3–10 gels (Serva) (Konarev et al., 2000a) and Phast gels pH 5–8 (Pharmacia), by thin layer gel-filtration (TLGF) in Sephadex G-50 (Konarev, 1982) or by SDS electrophoresis in Phast High Density gels after reduction with 2-mercapthoethanol. Proteins were stained with Coomassie BB R250. Horse myoglobin (pI 7.3) and cytochrome c (pI 10.6) were used as pI markers.

4.2.3. Detection of inhibitors

Proteinase inhibitors were detected after analytical separation of native proteins by the gelatin replicas method using a gelatin layer on photographic film (Konarev, 1986; Konarev et al., 2000a). Two or three replicas were sequentially placed in contact with the same IEF gel for 2, 5 and 20 min, respectively. The replicas were then applied to 0.8% (w/v) agarose gels containing 0.1 M Na_2HPO_4 (pH 9) and one of the following proteinases (Sigma); trypsin (1 μg/ml), chymotrypsin (10 μg/ml) and subtilisin (0.3 μg/ml), and incubated at 45 °C for 30 min. The activities of inhibitors in extracts were estimated based on the volume of extract required to obtain inhibitor banding patterns. Volumes below 0.3 μl, therefore, were defined as ''high activity''; volumes up to 2.5 μl as ''weak'' and the failure to detect digestion with 5 μl extract was defined as ''low or absent''.

4.2.4. Purification of proteins

Single inhibitor components were purified by combining affinity chromatography, reversed-phase HPLC (Konarev et al., 2000a) and micropreparative isoelectric focusing or SDS–PAGE. Ammonium acetate was added to the crude protein extract to 0.2 M and 1 l of extract

was shaken for 1 h with 15 ml of trypsin- or chymotrypsin–Sepharose gel. The gel was then washed with 0.4 l of 0.2 M ammonium acetate, 0.3 l of 0.1 M Na_2HCO_3 or 0.1 M Na_2HPO_4 and 0.3 l of water and packed into a column. Inhibitors were eluted with 0.015 M HCl, freeze-dried and subjected to HPLC on a C_{18} reversed-phase (RP) Vydac column with an acetonitrile gradient (ranges 20–50, 20–40 or 28–40% (v/v)). Fractions were freeze-dried and inhibitors detected by analytical IEF combined with the gelatin replicas method. HPLC-fractions containing proteinase inhibitors were separated on Servalyt precotes (pH 3–10) gels and transferred to PVDF-membrane or eluted from the gel. Protein bands were visualized by washing in a saturated solution of ammonium sulphate for 3 min, excised and eluted with 100–300 μl of water for 1 h. The eluate was then subjected to a second separation by RP–HPLC.

For micropreparative SDS–PAGE, the reduced protein fractions from HPLC were separated on a Phast high-density gel, transferred to PVDF-membrane by diffusion at 50 °C for 1 h and visualized with Coomassie BB R250.

Electrospray mass spectrometry was performed on a Thermoquest LCQ instrument operating in positive ion mode. Samples were dissolved in 1:1 methanol/water containing 1% (v/v) acetic acid and infused into the instrument at 3 μl/min. The spray voltage was 5kV and the capillary operated at 210 °C. Deconvolution was carried out using the Biomass program in Excalibar software version 1.2.

Cleavage of inhibitors by proteinases. Twenty micrograms of purified inhibitor dissolved in 100 μl of 0.1 M ammonium acetate, pH adjusted to 9.0 by ammonium hydroxide, was mixed with 20 μg of trypsin or subtilisin and incubated at 45° C for 4 days. Cysteine residues were blocked by *N*-isopropyl iodacetamide (NIPIA) after reduction with DTT.

Acknowledgements

IACR receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom. This work was carried out according to Program of Fundamental Investigations on Plant Protection of the Russian Academy of Agricultural Sciences and was supported under an "Ex-Agreement Visit" grant from Royal Society of London. The authors are grateful to Dr. L.I. Shashilova for providing plant accessions and to Dr. R. Fido for valuable advice and assistance.

References

Anisimova, I.N., 1996. Seed proteins as markers in *Helianthus*. In: Hind, D.J.N., Bentje, H.J. (Eds.), Compositae: Systematics, Proceedings of International Compositae Conference, Kew 1994, Vol. 2. Royal Botanic Gardens, Kew, pp. 627–641.

Ary, M.B., Shewry, P.R., Richardson, M., 1988. The amino acid sequence of a cereal Bowman-Birk type trypsin inhibitor from seeds of Job's tears (*Coix lachryma-jobi* L). FEBS Letters 229, 111–118.

Belozersky, M.A., Dunaevsky, Y.E., Musolyamov, A.X., Egorov, T.A., 1995. Complete amino acid sequence of the protease inhibitor from buckwheat seeds. FEBS Letters 371, 264–266.

Bhat, N.S., Puttarangappa, K., Virupakshappa, Prasad, D.T., 1996. Proteinase inhibitor in sunflower seed and its influence on growth and development of capitulum borer, *Helicoverpa armigera* (Hubner). In: Proceedings of 14th International Sunflower Conference, International Sunflower Association, Beijing-Shenyang, PR China, 12–20 June 1996, pp. 523–532.

Bremer, K., 1996. Major clades and grades of the Asteraceae. In: Hind, D.J.N., Bentje, H.J. (Eds.), Compositae: Systematics. Proceedings International Compositae Conference, Kew, Vol. 1, Royal Botanic Gardens, Kew, pp. 1–7.

Chen, L.J., Chen, P.W., 1997. EMBL Data Library. Accession No. U76004.

Cierpicki, T., Otlewski, J., 1999. Solution structure of *Linum usitatissinum* trypsin inhibitor. Unpublished. Direct Submission to Protein Data Base (08.12.1999).

Felizmenio-Quimio, M.E., Daly, N.L., Craik, D.J., 2001. Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from *Momordica cochinchinensis*. Journal of Biological Chemistry 276, 22875–22882.

Fujita, T., Kouchi, H., Ichikawa, T., Syono, K., 1993. Isolation and characterization of a cDNA that encodes a novel proteinase inhibitor I from a tobacco genetic tumor. Plant Cell Physiology 34, 137–142.

Kollipara, K.P., Hymowitz, T., 1992. Characterization of trypsin and chymotrypsin inhibitors in the wild perennial *Glycine* species. Journal of Agriculture and Food Chemistry 40, 2356–2363.

Konarev, Al.V., 1982. Component composition and genetic control of insect alpha-amylase inhibitors from wheat and *Aegilops* grain. Soviet Agricultural Science 6, 42–44.

Konarev, Al.V., 1986. Analysis of protease inhibitors from wheat grain by gelatine replicas method. Biochemistry (Moscow) 51, 195–201.

Konarev, Al.V., 1988. Polymorphism of chymotrypsin-subtilisin inhibitors and possibilities of its use in wheat variety identification. In: Biochemical Identification of Varieties. Materials III Int. Symp ISTA. Leningrad, USSR, 1987, Leningrad, VIR, pp. 176–181.

Konarev, Al.V., 1996. Interaction of insect digestive enzymes with plant protein inhibitors and host parasite co-evolution. Euphytica 92, 89–94.

Konarev, Al.V., Anisimova, I.N., Gavrilova, V.A., Rozhkova, V.T., Fido, R., Tatham, A.S., Shewry, P.R., 2000a. Novel proteinase inhibitors in seeds of sunflower (*Helianthus annuus* L.): polymorphism, inheritance and properties. Theoretical and Applied Genetics 100, 82–88.

Konarev, Al.V., Anisimova, I.N., Gavrilova, V.A., Shewry, P.R., 1999a. Polymorphism of inhibitors of hydrolytic enzymes present in cereal and sunflower seeds. In: Mugnozza, G.T.S., Porceddu, E., Pagnotta, M.A. (Eds.), Genetics and Breeding for Crop Quality and Resistance. Kluwer Academic Publishers, The Netherlands, pp. 135–144.

Konarev, Al.V., Anisimova, I.N., Gavrilova, V.A., Vachrusheva, T.E., Shashilova, L.I., Konechnaya, G.Yu., Shewry, P.R., 2000b. Heterogeneity and natural variability of proteinase inhibitors in sunflower (*Helianthus* L.) and other Compositae. In: Proceedings 15th International Sunflower Conference, 12–15 June, Toulouse, France, Vol. 1, Section A, pp. 109–114.

Konarev, Al.V., Anisimova, I.N., Vachrusheva, T.E., Konechnaya, G.Yu., Shewry, P.R., 2002a. Inhibitors of proteinases in safflower (*Carthamus* L.) and other Compositae. In: Proceedings International Safflower Conference, MT.

Konarev, Al.V., Kochetkov, V.V., Bailey, J.A., Shewry, P.R., 1999b. The detection of inhibitors of the *Sclerotinia sclerotiorum* (Lib) de

Bary extracellular proteinases in sunflower. Journal of Phytopathology 147, 105–108.

Konarev, Al.V., Tomooka, N., Ishimoto, M., Vaughan, D.A., 1999c. Variability of the inhibitors of serine cysteineproteinases insect alpha-amylases in *Vigna* and *Phaseolus*. In: Mugnozza, G.T.S., Porceddu, E., Pagnotta, M.A. (Eds.), Genetics and Breeding for Crop Quality and Resistance. Kluwer Academic Publishers, The Netherlands, pp. 173–181.

Konarev, Al.V., Tomooka, N., Vaughan, D.A., 2002b. Proteinase inhibitor polymorphism in the genus *Vigna* Savi subgenus *Ceratotropis* and its biosystematic implications. Euphytica, in press.

Kouzuma, Y., Kawano, K., Kimura, M., Yamasaki, N., Kadowaki, T., Yamamoto, K., 1996. Purification, characterization, and sequencing of two cysteine proteinase inhibitors, Sca and Scb, from sunflower (*Helianthus annuus*) seeds. Journal of Biochemistry (Tokyo) 119, 1106–1113.

Kumar, G.N., Houtz, R.L., Knowles, N.R., 1999. Age-induced protein modifications and increased proteolysis in potato seed-tubers. Plant Physiology 119, 89–100.

Levitskii, A.P., Pogoreletskaya, B.B. 1985. Proteinase inhibitors from sunflower seeds. Physiology and Biochemistry of Cultivated Plants (Kiev) 17, 153–157.

Li, Y., Huang, Q., Zhang, S., Liu, S., Chi, C., Tang, Y., 1994. Studies on an artificial trypsin inhibitor peptide derived from the mung bean trypsin inhibitor: chemical synthesis, refolding, and crystallographic analysis of its complex with trypsin. Journal of Biochemistry 116, 18–25.

Lin, X., Kaul, S., Rounsley, S., Shea, T.P., 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402 (6763), 761–768.

Luckett, S., Garcia, R.S., Barker, J.J., Konarev, Al.V., Shewry, P.R., Clarke, A.R., Brady, R.L., 1999. High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. Journal of Molecular Biology 290, 525–533.

Miura, S., Funatsu, G., 1995. Isolation and amino acid sequences of two trypsin inhibitors from the seeds of bitter gourd (*Momordica charantia*). Bioscience, Biotechnology and Biochemistry 59, 469–473.

Murray, C., Christeller, J.T., 1995. Purification of a trypsin inhibitor (PFTI) from pumpkin fruit phloem exudate and isolation of putative trypsin and chymotrypsin inhibitor cDNA clones. Biol. Chem. Hoppe-Seyler 376, 281–287.

Nagasue, A., Fukamachi, H., Ikenaga, H., Funatsu, G., 1988. The amino acid equence of barley rootlet trypsin inhibitor. Agricultural and Biological Chemistry 52, 1505–1514.

Okita, T.W., Cheesbrough, V., Reeves, C.D., 1985. Evolution and heterogeneity of the alpha-/beta-type and gamma-type gliadin DNA sequences. Journal of Biological Chemistry 260, 8203–8213.

Peterson, D.M., Forde, J., Williamson, M., Rohde, W., Kreis, M., 1991. Nucleotide sequence of a chymotrypsin inhibitor-2 gene of barley (*Hordeum vulgare* L). Plant Physiology 96, 1389–1390.

Richardson, M., 1974. Chymotryptic inhibitor I from potatoes. The amino acid sequence of subunit A. Biochemical Journal 137, 101–112.

Richardson, M., Cossins, L., 1974. Chymotryptic inhibitor I from potatoes: the amino acid sequences of subunits B, C, and D. FEBS Letters 45, 11–13.

Robert, L.S., Allard, S., Gerster, J.L., Cass, L., Simmonds, J., 1993. Isolation and characterization of a polygalacturonase gene highly expressed in *Brassica napus* pollen. Plant Molecular Biology 23, 1273–1278.

Ryan, C.A., 1990. Protease inhibitors in plants: gens for improving defenses against insects and pathogens. Annual Review of Phytopathology 28, 425–449.

Seemuller, U., Eulitz, M., Fritz, H., Strobl, A., 1980. Structure of the elastase-cathepsin G inhibitor of the leech *Hirudo medicinalis*. Hoppe-Seyler's Z. Physiol. Chem. 361, 1841–1846.

Shewry, P.R., 1999. Enzyme inhibitors of seeds: types and properties. In: Shewry, P.R., Casey, R. (Eds.), Seed Proteins. Kluwer Academic Publishers, The Netherlands, pp. 587–615.

Shewry, P.R., Lucas, J.A., 1997. Plant proteins that confer resistance to pests and pathogens. Advances in Botanical Research 26, 135–192.

Sossey-Alaoui, K., Serieys, H., Tersac, M., Lambert, P., Berville, A., 1994. Phylogenetic relationships of *Helianthus* species based upon RAPD fragments: amhiploid origin of the genus. In: Hind, D.J.N., Bentje, H.J. (Eds.), Compositae: Systematics. Proceedings International Compositae Conference. Royal Botanic Gardens, Kew, pp. 9–21.

Svendsen, I., Boisen, S., Hejgaard, J., 1982. Amino acid sequence of serine protease inhibitor CI-1 from barley. Homology with barley inhibitor CI-2, potato inhibitor I, and leech eglin. Carlsberg Research Communication 47, 45–53.

Valdes-Rodriguez, S., Segura-Nieto, M., Chagolla-Lopez, A., Verver, Y., Vargas-Cortina, A., Martinez-Gallardo, N., Blanco-Labra, A., 1993. Purification, characterization, and complete amino acid sequence of a trypsin inhibitor from amaranth (*Amaranthus hypochondriacus*) seeds. Plant Physiology 103, 1407–1412.

Williamson, M.S., Forde, J., Kreis, M., 1988. Molecular cloning of two isoinhibitor forms of chymotrypsin inhibitor 1 (CI-1) from barley endosperm and their expression in normal and mutant barleys. Plant Molecular Biology 10, 521–535.

Wilson, K.A., Chen, J.C., 1983. Amino acid sequence of mung bean trypsin inhibitor and its modified forms appearing during germination. Plant Physiology 71, 341–349.

Wingate, V.P., Broadway, R.M., Ryan, C.A., 1989. Isolation and characterization of a novel, developmentally regulated proteinase inhibitor I protein and cDNA from the fruit of a wild species of tomato. Journal of Biological Chemistry 264, 17734–17738.