1
Cupins: A New Superfamily of Functionally Diverse Proteins that Include Germins and Plant Storage Proteins

JIM M. DUNWELL

Department of Agricultural Botany, School of Plant Sciences, The University of Reading, Reading RG6 6AS, UK

Introduction

This review will describe the recent identification of a wheat protein, germin, whose function as an oxalate oxidase was first revealed in the early part of this century but which has remained largely overlooked since that time. Discovery of the sequence of this important enzyme (it is widely used in medical diagnosis) has now led to the further identification of a wide range of related proteins with a similar β-barrel structure. This new protein superfamily – designated 'cupins' – include several small microbial proteins, and many plant storage proteins that are duplicated versions of the underlying germin sequence. Most recently, homology modelling based on the structure of phaseolin has been used to generate a 3D structure of germin and it is expected that this information can be utilized to study the active site of this enzyme and its relatives.

Oxalate oxidase

HISTORICAL OVERVIEW

An enzyme with oxalate oxidase (ox–ox) activity, that converts oxalic acid to carbon dioxide and hydrogen peroxide, was first reported by Zaleski and Reinhard (1912) from studies of powdered wheat grains. Remarkably, it was more than 80 years before the identity of this enzyme was confirmed in cereals (Lane *et al.*, 1993; see below). After the first report, the enzyme was studied in both wheat flour and bran by Palladin and Lovchinovskaya (1916) and in 1928, Zaleski and Kukharkova concluded that the this enzyme should be considered as a dehydrogenase able to use only oxygen as a hydrogen acceptor. Meanwhile, Staehelin (1919) had reported a similar enzyme from leaves of several higher plants, including some known to contain oxalic acid, such as *Rumex*, *Rheum* and *Spinacia*. However, the most spectacular case of oxalic acid

Biotechnology and Genetic Engineering Reviews – Vol. 15, April 1998
0264–8725/98/15/1–32 $20.00 + $0.00

degradation was found in mosses – a source first discovered by Houget *et al.* (1927) in *Hypnum triquetrum*, *H. cupressiforme*, *Thudium abietinum* and *Polytrichum juniperum*. In subsequent work, Franke and Hasse (1937) characterized the enzyme from *Hyalocomium umbratum*, and noted a remarkable thermostability (cf. below for wheat germin). This was followed by Datta and Meeuse (1955) who designated the moss enzyme as a flavo-protein.

Amongst the cereals, ox–ox activity has probably been best characterized at a developmental level in sorghum, although in contrast to wheat, barley and rice (see below) no protein sequence information is available from this species. In the first study on sorghum, Pundir and Nath (1984) used the 15 000 *g* supernatant from 10-day-old leaves and found enzyme activity with a pH optimum of 5 and an absolute requirement for metal ions, particularly iron. This finding seems to be contradicted by the conclusion from subsequent studies (Pundir, 1991; Kuchhal *et al.*, 1993) that copper is the critical metal ion. This latter conclusion was supported by a study on roots from the same material (Pundir and Kuchnal, 1989) which showed an involvement of copper in promoting enzyme activity.

Similar studies on barley ox–ox have been conducted by Chiriboga (1966) and Sugiura *et al.* (1979). The most recent, and most thorough, study of the barley enzyme is that reported by Kotsira and Clonis (1997) who purified ox–ox to homogeneity by thermal treatment (60°C for 10 min), followed by two affinity chromatography steps. It was concluded that the enzyme is activated by 1 mM each of Ca^{2+} and Pb^{2+}, and that some of the cysteines may act to preserve the catalytic activity by maintaining the integrity of the pentameric assembly via disulphide bond formation. Optimization of the assay for ox–ox activity has been reported by Zhang *et al.* (1996) who showed a 10–100-fold increase in sensitivity if 60% ethanol is included in the assay mixture.

COMMERCIAL SIGNIFICANCE OF OXALATE OXIDASE

The commercial significance of this enzyme comes largely from its widespread use in kits to assay oxalate levels in blood plasma and urine. Indeed, regular assessment of such levels is required for control of hyperoxaluria, whether it be the rare genetic condition which leads to deposition of calcium oxalate throughout the body (De Pauw and Toussaint, 1996), or the much more frequent incidence of urolithiasis – the formation of kidney or bladder stones (Ruml *et al.*, 1997).

By far the most common source of enzyme used in these kits (e.g. Sigma or Boehringer Mannheim) is barley roots. Although the assays are simple and quick, continuous efforts are being made to improve the accuracy still further. Such studies include immobilization of the enzyme either on glass beads (Hansen *et al.*, 1994), on AF-Tresyl Toyopearl 650 gel (Yamato *et al.*, 1994), on a pre-activated nylon membrane (Saka Amini and Vallon, 1994), or by addition of ethanol to the assay mixture (Zhang *et al.*, 1996). The products can be quantified by colorimetric (Petrarulo *et al.*, 1994), amperometric (Saka Amini and Vallon, 1994; Yamato *et al.*, 1994) or chemiluminescence (Balion and Thibert, 1994) methods. Similar reports on the enzyme extracted from beet stems (Obzansky and Richardson, 1983,1984; Varalakshmi *et al.*, 1995) include one on immobilization on concanavalin A (Varalakshmi and Richardson, 1992), a method claimed to eliminate the significant inhibition produced by azide, nitrate (Meeuse and Campbell, 1959) and glycolate.

Alternative sources of enzyme include moss (Lakeret *al.*, 1980) and especially banana peel – the preferred source in tropical countries (Inamdaret *al.*, 1986,1989; Lathika *et al.*, 1995a, b). Degradation of the oxalate found in bananas *in situ* had been known for many years previously (Wyman and Palmer, 1964; Shimokawa *et al.*, 1972).

OXALATE AND HUMAN METABOLISM

Although a high level of oxalate in plasma is usually considered to be a pathological condition (see previous section), remarkably high values in blood cells, with a maximum of 2910 µmoles/l, were reported by Albrecht *et al.* (1994). These authors suggest that oxalate may not simply be an end product of metabolism, but rather it may have an important physiological function, perhaps by providing the hydrogen peroxide generated by phagocytes during the inflammatory process.

OXALATE DEGRADATION AND GENE THERAPY

As described briefly above, primary hyperoxaluria is a rare, and sometimes fatal, hereditary disease (De Pauw and Toussaint, 1996). There are known to be two forms, type 1 which is due to a deficiency of the liver-specific peroxisomal enzyme alanine:glyoxylate aminotransferase/serine pyruvate amino-transferase, and type 2 which concerns glyoxylate reductase/D-glycerate dehydrogenase, a cytosolic enzyme present in leucocytes and hepatocytes. Several gene therapy approaches are being considered to alleviate these conditions. They range from introduction of the appropriate wild-type version of the endogenous gene to expression of an alternative oxalate degrading enzyme. The latter approach was presaged by the observation that rats implanted with dialysis membrane capsules containing immobilized ox–ox from banana (see above) could effectively metabolize intraperitoneally injected [^{14}C] oxalate as well as its precursor, [^{14}C] glyoxylate (Raghavan and Tarachand, 1986) (see also Batich and Vaghefi, 1994).

Although their eventual ambition was to clone the fungal oxalate decarboxylase gene for use in gene therapy (Finlayson and Peck, 1988; Sidhu *et al.*, 1997) Lung *et al.* (1991,1994) did succeed in cloning a bacterial equivalent, the oxalyl-CoA decarboxylase from *Oxalobacter formigenes*, an oxalate degrading anaerobic bacterium found in the mammalian intestine (Baetz and Allison, 1989; Cornick and Allison, 1996; Kuhner *et al.*, 1996). This approach has recently been superseded by the cloning and patenting of the human version of this enzyme (Coleman *et al.*, 1997) and the publication of several related expressed sequence tags (ESTs) from human (gil2238232, gil1727767, gil1722162), mouse (gil2192248), Caenorhabditis (gil1044814) and yeast (gil602387) sources. It is likely that the recent advances in understanding the sequence and structure of oxalate degrading enzymes (see below) will help in this area of research, which in many ways is equivalent to the work on transgenic plants also described below.

OXALATE DEGRADATION AND BIOREMEDIATION

An interesting, non-medical application of oxalate degradation is provided by the example (Anon., 1992; Morton *et al.*, 1994) of using *Pseudomonas oxalaticus*

(Chandra and Shethna, 1975), a bacterium found in rhubarb patches, to digest sodium oxalate, a hazardous byproduct of the Bayer process (Farquharson *et al.*, 1995) utilized in aluminium smelters. The specific bacterial enzyme responsible for this degradation has not been isolated to date.

Germins and germin-like proteins from plants

GERMIN IN CEREALS

Discovery of oxalate oxidase activity

The onset of growth in germinated wheat embryos is accompanied by the production of a soluble, pepsin-resistant homopentameric glycoprotein (oligomer approx. 125 kD) that was initially called 'g' and then named 'germin' (Grzelczak and Lane, 1984; McCubbin *et al.*, 1987). In subsequent studies, the sequence of two very similar germin genes designated *gf-2.8* and *gf-3.8* were obtained by Dratewka-Kos *et al.* (1989) and Lane *et al.* (1991). The isoforms of germin are discrete temporal markers of wheat embryo development; wall-associated germin accounts for about 40% of total germin in germinating wheat embryos, with the appearance of germin in the apoplast being the most conspicuous germination-related change in the distribution of cell-wall proteins (Lane *et al.*, 1986,1992). Although these studies revealed much about details of the germin protein and its expression, its function remained unknown for several years. Only when the sequence of barley ox–ox became available was it revealed that germin gf-2.8 was in fact a protein with ox–ox activity (Lane *et al.*,1993; Dumas *et al.*, 1993) and was therefore presumably the enzyme first identified some 80 years previously (Zaleski and Reinhard, 1912). It is thought that the closely related gf-3.8 is monomeric and does not exhibit ox–ox activity. This difference creates a potential problem of nomenclature, in that similar genes encode proteins with quite different properties, and therefore functions. It is proposed to restrict the term germin to the cereal genes and to stress that the gene products may have a range of function.

Germins in cereals

It was discovered at an early stage in the study of germin that an immunologically similar protein could be detected in all cereals tested, i.e. in the stems of barley, oat, and rye (Grzelczak *et al.*, 1985) and in maize and rice (Lane *et al.*, 1991). However, no such protein was detected in pea, lettuce or radish, nor in the moss *Tortula ruralis*, although as the authors state, 'such findings do not bear, definitively, on the occurrence of proteins related to germin' (see below).

Of all cereals tested to date, the largest number of germin-like gene sequences have been found in rice. These comprise one 20 AA fragment (gil348652) from a 25 kD protein and 20 ESTs. The EST sequences are from 8-day-old shoots, either green or etiolated, or from the panicle at the ripening stage (*Table 1*). Translation and manual editing of obvious sequencing errors reveals a possible total of eight distinct proteins which can be divided into three classes (Dunwell, unpublished results). The first is most similar (approx. 70%) to the Prunus ABP 20 protein (gil1916809), the second is related to the barley X93171 protein, and the third is similar to the Arabidopsis GLP

Table 1. List of Rice ESTs which encode germins

Number	dbEST Id	Library
1	71171	shoot
2	71604	shoot
3	71916	shoot
4	72151	shoot
5	72154	shoot
6	72327	shoot
7	72360	shoot
8	72415	shoot
9	75576	shoot
10	75954	shoot
11	144047	green shoot
12	144301	green shoot
13	144848	green shoot
14	145340	green shoot
15	146125	green shoot
16	146169	green shoot
17	146752	green shoot
18	146941	green shoot
19	726312	panicle at ripening stage
20	726835	panicle at ripening stage

2b germin-like protein. Interestingly, there are no examples closely related to the wheat germin or barley root ox–ox.

Patterns of expression

In 3-day-old barley seedlings, oxalate oxidase is localized in the epidermal cells of the mature region of the primary roots, and in the coleorhiza, whereas after 10 days of growth it is only detected in the coleorhiza (Dumas *et al.*, 1995).

Function

Development. Two features of wheat germin have been linked to specific roles in plant development (Lane, 1991), particularly in the modification of the extracellular matrix (Lane, 1994). First, the tight adhesion of germin to the highly substituted glucuronogalactoarabinoxylans (GGAX) suggests a possible role in incorporation of these compounds into the growing wall. Secondly, the generation of H_2O_2 is required for peroxidase-mediated reactions such as lignification, and the cross-linking of coumarates, extensins and ferulates to cell wall hemicelluloses and pectins. In direct contrast to the first role, these reactions would stop the process of wall extension. The presence of several auxin responsive elements in the promoter of germin *gf-2.8* (Berna and Bernier, 1997) is additional circumstantial evidence for a link between cell wall pH and germin activity (Lane *et al.*, 1992).

Plant pathogen interactions. In a preliminary study of defence-related gene

expression in barley, Wei (1994) identified a cDNA clone (gi|1070358, X93171) with approximately 45% identity to barley root oxalate oxidase. It is now known to be most similar (90%) in sequence to a reconstructed protein encoded by the fusion of two rice ESTs, gi|1632028 and gi|1632551. Subsequently, there have been two particularly significant publications (Dumas *et al.*, 1995; Zhang *et al.*, 1995) describing the involvement of ox–ox in the response of barley plants to powdery mildew infections. In the former study, it was shown by immunoblots and direct assays that the enzyme was induced in leaves between three and five days after inoculation, and that it was localized mainly along the vascular bundles. There was no induction by wounding. The latter investigation demonstrated a much more rapid response, with colorimetric assays and activity blots showing a detectable increase after 24 h and a 10-fold increase after 48 h. These increases appeared 1–3 days earlier than PR-1 accumulation and it is suggested that the activity results from *de novo* protein synthesis, though pathogen-induced enzyme activation is also possible. It is not known whether the H_2O_2 generated by the enzymatic process is used as a signal for the induction of other components of the defence response, for lignification, or directly to inhibit pathogen development (Thordahl-Christensen *et al.*, 1997; Wojtaszek, 1997).

Similar results were reported by Hurkman and Tanaka (1996b) for wheat leaves inoculated with the same pathogen.

Salt stress. There have been several, related studies on the effects of salt stress on the expression of germin genes in barley (*Hordeum vulgare* L. cv CM 72) seedlings (Hurkman *et al.*, 1989, 1991, 1994; Hurkman and Tanaka, 1996a). In control seedlings, germin mRNA levels (as estimated by expression of sequence P45851) are regulated developmentally, with the highest amount in roots and the lowest in shoots. These levels are maximal 2 days after inbibition, before declining. In contrast, for seedlings grown in the presence of 200 mM salt, the levels remain high for an additional day. Expression patterns in salt stressed roots are more complicated, with levels rising only transiently when 6-day-old roots are exposed to 200 mM NaCl (Hurkman and Tanaka, 1996a).

OXALATE CONTENT OF CEREALS

In the light of the discoveries that a predominant germination-related protein in cereals has ox–ox activity and that this protein has several close relatives, it is pertinent to consider the state of knowledge of oxalate in cereals, particularly since this compound has never been considered to have any significance in this group of plants. For a general review of oxalate in crop plants, the reader is referred to the survey by Libert and Franceschi (1987).

There have been few efforts to localize oxalate in cereal plants. In an interesting and possibly unique histological study, Prankerd (1920) reported the presence of calcium oxalate crystals in the nodes of the wheat plant. It was inferred that these crystals played a role in the detection of gravity and they were estimated to travel from one side of the cell to the other in under five minutes. Much later, Wagner (1981) reported that oxalic acid in barley is mostly localized within the vacuoles.

Attempts to quantify the levels of oxalate in cereals are also scarce. Nelson and Hasselbring (1931) estimated the amount in wheat plants to be 0.02% on a fresh

weight basis (0.11% by dry weight), and values of 0.019% for barley, 0.029% for maize, 0.04% for oats, and 0.048% for rye were calculated by Nelson and Mottern (1931). In a more detailed physiological study on maize, Wadleigh and Shive (1939) measured values of oxalic acid on a dry weight basis, ranging from 0.062% for plants grown at pH 3.0 to 0.178% for plants at pH 8.0. However, in the same year, Olsen (1939) could not detect oxalate in leaves of barley and maize, whereas in a study of nutrient conditions on the growth of oats, Pepkowitz *et al.* (1941) showed the highest oxalate levels in those plants supplied with additional nitrate. The most relevant of the studies from that period is probably that of Andrews and Viser (1951) who measured the oxalate content of wheat embryos and found a value of about 1 μg/embryo. Any overall conclusion from these studies must be qualified by the proviso that the accuracy of the values cannot be assessed according to modern standards, but nevertheless they do show that low levels of oxalate are recorded in most cereals, and that levels in whole plants are modified by a range of environmental variables.

A specific defence role for oxalic acid in rice was suggested by the findings of Yoshihara *et al.* (1990). These authors showed conclusively that this acid could be isolated from the leaf sheath and that it acted as a potent inhibitor of sucking by the brown planthopper.

TRANSGENIC PLANTS EXPRESSING OXALATE OXIDASE

To date, the only practical application of germins in transgenic crops is based their activity as an ox–ox, and its potential use as an antifungal agent.

Oxalic acid is produced by several plant pathogenic fungi (Loewus *et al.*, 1995) and is considered to play a major role in the pathogenicity of *Sclerotinia sclerotiorum* (Noyes and Hancock, 1981; Rowe, 1993), a particularly important pathogen of sunflower (Masirevic and Gulya, 1992; Sackston, 1992) and many other dicot crops. Evidence for the role of oxalic acid comes from studies on mutant strains of the fungus which are deficient in oxalate production and also avirulent. Revertants for acid production regain their virulent nature (Godoy *et al.*, 1990).

The proposed mode of action for oxalic acid in producing disease symptoms is first the chelation of calcium from the pectate fraction of the xylem and associated pit vessels, secondly the entry of air leading to xylem embolisms and wilting (Sperry and Tyree, 1988), and finally the reduction in pH which stimulates the activity of the fungal enzymes such as polygalacturonase, methyl esterase and cellulase (Marciano *et al.*, 1983).

Analysis of the role of oxalic acid, as described above, led to the strategy of introducing oxalate degrading enzymes into plants as a means of protecting them against the fungal toxin (Thompson *et al.*, 1995a,b). This transgenic strategy, which has some parallels with the work on human gene therapy mentioned above, has utilized both available genes, the ox–ox from barley or wheat (see above) and the fungal oxalate decarboxylase (see below). Field studies are most advanced on transgenic oilseed rape material containing the former gene (J.M. Dunwell and C. Thompson, unpublished results). When grown in field conditions containing a high level of *Sclerotinia*, the transgenic plants show a significant level of protection against the pathogen. It is not yet known whether this is sufficient for commercial exploitation.

A similar transgenic approach, utilising bacterial genes for oxalate degradation (cf.

Koyama, 1988), has been suggested in a study of *Sclerotinia* infection of *Arabidopsis* (Dickman and Mitra, 1992).

In addition to their value in applied studies described in the previous sections, transgenic plants have also been used in a series of more fundamental studies such as that in which the wheat *gf-2.8* germin gene was introduced either as the intact gene, or as promoter-*GUS* fusions, into transgenic tobacco (Berna and Bernier, 1997). Heterologous gene expression was monitored *in vitro* and *in vivo* by GUS or ox–ox activity, and shown to occur in developing seeds and in seedlings. This transcription was stimulated by auxin, probably owing to the presence of several auxin-responsive elements (e.g. TGTCCCAT and TGTCTC) in the promoter sequence. Interestingly (in view of the germin/vicilin connection; see below), this sequence also contains nine 'legumin boxes' scattered between −1530 and −1142 upstream. Analysis of the transgenic protein itself revealed two enzymatically active isoforms, corresponding to the wheat germins G and G′ which differ by the presence of antennary N-acetylglucosamines (Jaikaran *et al.*, 1990). This analysis suggests that the post-translational modifications of the protein (oligomer assembly, stability, glycosylation) are very similar, if not identical, in the homologous and heterologous hosts.

GERMIN-LIKE PROTEINS (GLPs)

This group of proteins includes all those plant proteins closely related to the cereal germins. It is, therefore, a somewhat arbitrary designation.

Gymnosperms

Pinus caribea. In a study of extracellular glycoproteins secreted by embryogenic cultures of the Caribbean pine (*Pinus caribea* Morelet var *hondurensis*), Domon *et al.* (1995) identified three proteins with some similarity to germins. This similarity was based on N-terminal sequence from one protein, and cross reactivity with antibodies raised to the nonglycosylated wheat monomer as well as to the glycosylated barley pentameric protein. Interestingly, these proteins were absent from the control non-embryogenic cultures and they therefore represent the first markers of somatic embryogenesis in this species.

Pinus radiata. A study of a cDNA library made from somatic embryogenic tissue has recently revealed four ESTs (gil1839644, 1839645, 1839658, and 1839659) with significant similarity to germins (S.L. Bishop-Hurley *et al.* unpublished results); Northern blotting using 5 μg RNA showed no expression in roots, stems, needles and callus. Manual editing (J.M. Dunwell, unpublished results) on a combination of the translated products from these ESTs suggest that they probably represent a gene whose protein product is given in Table 2. This sequence is most similar (44% identical, 62% similar) to the Arabidopsis germin-like protein GLP8 (gil1755192).

Dicot species

Arabidopsis. In Arabidopsis, there are a total of 27 protein sequences (gil1755152-

Table 2. Putative sequence of germin-like proteins from *Pinus radiata* (a) and Strawberry (b). These sequences have been produced by translation and manual editing of various ESTs (J.M. Dunwell, unpublished results, see text for details)

(a)
DPEALQDFCVADTKSSIFMNGRPCLNPMEASPEHFMTSALRKAGNLSANPLGFSVIVTTPANMPGLNTLGLSMG
RIDMAPGSAIPPHIHPRGSETIXVVRGALNVGFVDTSNRLXXHKLVTGDVFVFPKGAVHYLQNIGKITAFIVSA
FNSQNPGTAIVSLATFASNPAIPYEVLSKAFAISVQEVSQIRKSFGGT

(b)
SYAAMQDFCVADYAAPQSPAGYACKDPAKVTVDDFVFSGLRVAGNISSINKIGLSAAFAVNFPGLNGLGVALVR
ADFAVGGVVPIHSHRDATELVILVERTVVAGFIAXNNKALCEXSAQGGDTMVFPQGLFHFLMNVGRTPALXYAS
FSSANPGVQTLEXTLFXNELPTEIIXNSTLLEKXXIXKL

1755192, gil1934726-1934730, gil1107491, gil1592672, gil2239042), two BAC sequences (gil2091318,gil2092889 which encode in the negative strand a protein very similar to gil1755165), and 31 ESTs in the databases (see also Delseny *et al.*, 1994; Hofte *et al.*, 1993) and these can themselves be sorted into a number of classes (J.M. Dunwell, unpublished results). The proteins seem to comprise 12 distinct sequences, reduced to 8 if a more conservative estimate is made.

Brassica. The only full length sequence of a Brassica germin-like gene in the GenBank database is accession U21742 produced from 5-day-old etiolated seedlings of *Brassica napus* cv Samourai. Its closest neighbour (87% identical, 92% similar) is an Arabidopsis 'germin' GLP 3 (U75189). However, three other partial sequences have been isolated from *B. napus* by the use of PCR based methods (A. Walker and I.J. Evans, unpublished results) and these are described in *Table 3*. In terms of similarity, clones 1 and 2 are most similar (84%) to a section of Arabidopsis GLP6 (U75194) and clone 3 shows close similarity (73%) to a section from the barley germin X93171. Other ESTs are also known, gblH07760 from *B. napus* and gblL46495 from *B. campestris*. With the insertion of an additional 't' at position 174, the former EST encodes a sequence identical, with the exception of one amino acid, to the *B. napus* sequence U21742 (J.M. Dunwell, unpublished results).

Sinapis. As part of an attempt to analyse rhythmic phenomena in the long day plant *Sinapis alba*, Heintzen *et al.* (1994) searched for mRNAs whose concentration varied with the time of day. They identified a germin-like transcript that underwent circadian oscillations in light/dark cycles with maxima at around 8–9pm and minima at around 8am. At a cellular level, the transcripts were expressed in the epidermis and spongy parenchyma of young leaves, and in distinct regions of the epidermis and cortex of stems and petioles. The protein itself (gil683488) was associated with the primary cell walls.

Pharbitis. In a very similar study to that on *Sinapis* described above, Ono *et al.* (1993) used the short day plant *Pharbitis nil* in an examination of floral induction. The convenience of this species is clearly demonstrated by the fact that a single 16-hour dark treatment is sufficient to induce flowering. During the second half of the inductive dark period the level of a specific 22 kD protein (isoelectric point 7.5)

Table 3. Germin-like sequences from *Brassica napus* (A. Walker and I.J. Evans, unpublished results)

Clone 1
CGTCAGTCCCCCGCCGCATATGCATCCGCGGGCGACTGAGATCTTGGTTGTCCAACAAGGAACCTTGCTTGTAGG
GTTTGTGTCTTCAAACCAAGATGAGAACCGTCTTTTCGCCAAAACGCTCAATGTTGGTGATGTATTTGTGTTTC
CAGAAGGCTTCCTCCATTTTCAACTCAAGGCATCCGGGCTAGAGCGGCGCACCGCGGTGGAGCTCCCAG

Translation:
RQSPPHMHPRATEILVVQQGTLLVGFVSSNQDENRLFAKTLNVGDVFVFPEGFLHFQLKASGLERRTAVELP

Clone 2
CCGCCGCATATGCATCCGCGGGCCACTGAGATCTTGATTGTCCAACAGGGAACCTTGCTTGTAGGGTTTGTCTC
TTCAAACCAAGATGGAAACCGTCTTTTCGCCAAAACGCTCAATGTTGGTGATGTATTTGTGTTTCCAGAAGGCC
TCATCCATTTTCAACTAAGGCATCCAGGGCTAGAGCGGCGCACCGCGGTGGAGCTCC

Translation:
PPHMHPRATEILIVQQGTLLVGFVSSNQDGNRLFAKTLNVGDVFVFPEGLIHFQLRHPGLERRTAVEL

Clone 3
Translation (by manual editing):
PPHMHPRAREILTVLEGTLRVGFVTSNPNNRFTKVLKKGDVFVFPVGLVTFQSNVGNGNAVAIAALSSQNPGA
ITIANXVFG

increased significantly. This protein (gi|662292) was subsequently identified (Ono *et al.*, 1996) as a germin-like protein whose mRNA was detected particularly in the young expanded cotyledon and leaf. The level peaked about 10 hours from the beginning of the dark period, in contrast to *Sinapis* where the peak for the related protein occurs about 12 hours after the start of the light period (Heintzen *et al.*, 1994).

Interestingly, this Pharbitis protein was tested and had no ox–ox activity.

Mesembryanthemum. *Mesembryanthemum crystallinum* is a facultative halophyte and has been widely used in studies of biochemical changes and gene expression during water stress (Bohnert *et al.*, 1988). In a detailed biochemical study, Adams *et al.* (1992) found an increase in oxalate content in the bladder cells from <1 mM up to 106 mM as salt levels were increased from 1 mM to 5 mM. This finding may be linked to the results obtained from a differential screening of a cDNA library produced from roots subject to 6 hours of salt stress (500 mM) (Michalowski and Bohnert, 1992). These authors identified a significant decline in the transcript of a germin-like protein (gi|1169942). In a subsequent study of changes in transcript levels for several genes, Andolfatto *et al.* (1994) found that a reduction of 'germin' message in roots after 30 hours of salt stress (400 mM) but an increase after 120 hours.

Detailed comparison of the protein sequence from this species with other germin-like proteins reveals an obvious loss in similarity towards the C-terminus at amino acid position 197. Before that residue, there is approximately 75% similarity to Arabidopsis germin U75192. Beyond that residue there is no significant similarity to any other germin. However, if an additional nucleotide is inserted at that point (position 609 in gi|167257), then the similarity is regained (J.M. Dunwell, unpublished results). This suggests either a sequencing error which introduced the effect of a frame shift mutation, or that indeed such a mutation occurred during the evolution of this protein. There is no evidence available for distinguishing between these alternatives.

Prunus. There have been two recent database submissions of germin-like sequences (U81162,U81163) from peach (*Prunus persica* cv Akatsuki). The authors of this study describe these sequences as auxin-binding proteins (ABP19/20) (see below for discussion of ABPs), but this is an obvious misnomer since their closest relative in the public database is an Arabidopsis 'germin' sequence GLP 2B (U75193). However, they are even more similar to a previously unpublished sequence from strawberry (see below), another member of the Rosaceae.

Strawberry. As part of a large-scale EST project on gene expression in strawberry fruit, a germin-like sequence has been identified recently (A. Van Tunen, personal communication) (see *Table 2*). Not surprisingly, it is most similar (54% identical, 72% similar) to the sequence from peach mentioned in the previous paragraph.

Legumes. Although Staehelin (1919) reported the oxidation of oxalic acid by flour produced from a range of legumes, namely *Lupinus albus, Phaseolus vulgaris, Glycine soya, Vicia faba*, and *Pisum sativum*, the first evidence for the presence of a germin-like protein in any legume, came from the work of Swart *et al.* (1994) on a cell wall protein from pea roots. This protein was implicated in the attachment of *Rhizobium* and *Agrobacterium* bacteria to the plant cell wall, a process mediated by the bacterial protein rhizadhesin. Analysis showed the putative receptor to be a glycoprotein with an isoelectric point of 6.4, a molecular mass of 32 kD before, and 29 kD after glycosidase treatment, and to have an N-terminal sequence of ADADALQDLCVADYASVILVNGFASKPLI (with some doubt over the last three residues). Use of this sequence in a BlastP search reveals the closest relative (69% identity) to be an Arabidopsis germin (gi|1934730) and two ESTs (W43342, AA042591). It may be of interest to note that the most successful extraction method for removing the receptor from the cell wall involved the use of an oxalate/oxalic acid mixture.

More recently, this evidence from pea has been supplemented by the direct measurement of oxalate oxidase activity in nodules of faba beans (*Vicia faba* L.) (Trinchant and Rigaud, 1996). These authors, who had previously detected high levels (70 mM) of oxalate in nodules (Trinchant *et al.*, 1994), have now demonstrated that application of a water stress increases the level of oxalate oxidase activity fourfold, and reduces the level of oxalate by 55%. The enzyme itself is characterized by an optimal activity at pH 8, in contrast to the cereal enzyme referred to above, which has an optimum at pH 3. It is suggested that the oxalate present at high levels in the nodules could play a role as a complementary substrate for bacteroids, and as a means to slow down the decline in nitrogen fixation induced by water restricted conditions.

In the light of these findings of high levels of oxalate in nodules it is somewhat surprising that legumes show susceptibility to *Sclerotinia* (see above).

MAPPING INFORMATION FOR GERMINS AND GERMIN-LIKE PROTEINS

There have been two studies to locate the position of the germins on the cereal genome map. In the first, quoted in Lane *et al.* (1991) it was reported that a germin cDNA used as a probe mapped primarily to chromosomes 4A (~5 copies), 4B (~3 copies) and 4D (~9 copies) in hexaploid wheat. In the second, Dubcovsky *et al.* (1995) used the barley

clone pWJHGermin (Hurkman *et al.*, 1994) on a mapping population of *Triticum monococcum* and found one pair of fragments mapped to the centromeric region of chromosome $2A^m$ (arm location unknown) and another locus was mapped to chromosome $4A^m$. The position of this locus is similar to that previously reported by Lagudah *et al.* (1991) for *Triticum tauschii* and Devos *et al.* (1995) for wheat chromosomes 4B and 4D. In an unpublished study (Gale *et al.*) quoted in Berna and Bernier (1997), it has been shown that the promoter of the *gf-2.8* gene is largely confined to the D genome of wheat; it hybridizes to four genomic fragments from this genome, one from the A genome and appears to be absent from the B genome.

The only information from rice is unpublished data provided by Dr Ilkka Havukkala (Rice Genome Research Program) who reports that the rice clone S1623 (=D39942) maps to the top terminus of chromosome 4 (Koike *et al.*, 1997), at the same location as marker C445 on the map published by Kurata *et al.* (1994). It is likely that there is only one location for this clone in the genome.

There is no published information on the chromosomal location of any of the dicot GLPs, although it can be expected that the Arabidopsis genes will be placed on the map in the near future.

Microbial and animal proteins with sequence similarity to germins

SPHERULINS AND DESIGNATION OF THE 'GERMIN BOX'

The first proteins found to have any significant sequence similarity to the wheat germins were the spherulins, proteins expressed in the slime mould *Physarum polycephalum* during the process of plasmodial encystment or spherulation (Bernier *et al.*, 1987). Plasmodia will encyst to form spherules when subjected, in the dark, to dehydration, cooling, acidic pH, sublethal concentrations of heavy metals, high concentrations of some carbohydrates, or starvation. These spherules are hard-walled oligonucleated subunits produced by cleavage of the plasmodia, and they can remain viable for many years.

At the amino acid level there is 44% similarity in sequence between germin gf-2.8 and spherulin 2b (Lane *et al.*, 1991), a value that increases to about 60% for the central core sequence. Near the centre of this conserved core is the decapeptide sequence PH(I/T)HPRATEI, a region designated as the 'germin box'. In the sections below, it will be demonstrated that this motif is conserved across a much wider range of proteins and is part of the putative active site of the cereal ox–ox enzymes (Gane *et al.*, 1998). It should be noted however that the spherulins do not show ox–ox activity and their biochemical function remains unknown.

IDENTIFICATION OF TWO HISTIDINE-CONTAINING MOTIFS

In a search for possible germin progenitors, BLAST searches were conducted on various 'germin' and 'germin-like' sequences in the GenBank database. Subsequent analyses (J.M. Dunwell, unpublished results) using the BLOCKS (Henikoff and Henikoff, 1991), MEME (Bailey and Elkan, 1994) and PROSITE programmes showed first that only certain residues within the 'germin box' are absolutely conserved across all proteins, and secondly that other regions of the protein are equally

conserved (Dunwell and Gane, 1998). Specifically, it was shown that the 'germin box' is part of a longer 20/21 AA motif, $G(X)_5HXH(X)_{11}G$. This sequence is part of the PRINTS motif 'GERMIN1' and is followed (usually after 15 residues) by a second motif of 16 AAs, $G(X)_5P(X)_4H(X)_3N$ (part of PRINTS motif 'GERMIN2'). Using the bean storage protein phaseolin (Lawrence *et al.*, 1994) as a reference, these two histidine-containing motifs are part of the β-strands designated C/D and G/H within the two β-barrel elements. Considering the first motif, the two flanking glycines are equivalent to the strictly conserved Gly^{249} and Gly^{269} which assist in the formation of the short interstrand loops B–C and D–E within the C-terminal β-barrel of phaseolin. Similarly, the second motif contains a proline equivalent to conserved Pro^{303} which is part of the interstrand loop between strands G and H. The variable spacing between motifs is equivalent to the insertions (varying between 7 and 25 residues) tolerated in the E–F loop (see Gane *et al.*, 1997).

CUPINS

The recent sequence analysis reported by Dunwell and Gane (1998) identified a series of protein families (many bacterial) which contain the two motifs described in the previous paragraph. These families are considered below and have been divided into two categories. The first includes the proteins having a single 'germin-type' domain, and the second include those with a duplication of this domain. Amongst the list of single domain proteins are several that are much smaller than the germins, with the most extreme being only 77 residues in length. These 'protogermins' have only a core β-barrel structure (as judged by homology analysis), with a minimal intermotif spacing and without any terminal α-helices. It is proposed to designate them as 'cupins' (from the Latin term 'cupa' for a small barrel or cask). Whilst it is acknowledged that enzyme nomenclature is usually based on function, in this case the related nature of the proteins (many with no known function) has been determined on the basis of sequence/structure. Using this nomenclature, the 'germins' and 'germin-like' proteins (mean length *c.* 200 AA) may be considered as 'long cupins', and the oxalate decarboxylase and other two-domain proteins (see below) as 'bi-cupins'.

ONE-DOMAIN PROTEINS

Auxin-binding proteins (ABPs)

These proteins are widespread in the plant kingdom, and are thought to act as a receptor for IAA, and thereby control a range of plant growth responses (Millner, 1995; Napier, 1995; Napier and Venis, 1995; Venis and Napier, 1995). Amongst several descriptions of such proteins from both monocots (Palme *et al.*, 1992; Schwob *et al.*, 1993) and dicots (Choi, 1996; Leblanc *et al.*, 1997; Lazarus and Macdonald, 1996; Shimomura *et al.*, 1993) is one on the isolation of a presumed ABP from peach shoot apices (Ohmiya *et al.*, 1993). However, this protein binds 2,4-D with low affinity (Kd=40μM), it has even lower affinity to IAA and NAA, and it does not cross-react with antibodies to maize ABP1 (Fellner *et al.*, 1996). As Jones (1994) states, its function remains unknown. Recently, the sequences of two peach proteins gi|1916807 and gi|1916809 (Accessions U81162,U81163), described as ABPs, were submitted to

the GenBank database (A. Ohmiya *et al.*, unpublished results). As mentioned above, this is presumably a misnomer, since they show much closer sequence similarity to germins than to functional ABPs (see details above). This misinterpretation can now be explained by the fact that ABPs belong to the one-domain cupin superfamily (Dunwell and Gane, 1998).

These authors recognized a consistent similarity between the sequence of ABPs and that of germins, particularly in the central core sequence of three ABP exons that include the 'germin box' (motif 1). Within the ABPs, this motif is known as the 'D16' sequence (Venis and Napier, 1995) or 'Box A', and is thought to be involved in the binding of the carboxylic acid group of IAA (Brown and Jones, 1994; Edgerton *et al.*, 1994; Jones, 1994). Further evidence suggesting a link between ABP1 and germins come from circular dichroism spectral analyses and structure predictive algorithms, neither of which support a significant amount of helical secondary structure (Jones, 1994). This is confirmed by the molecular modelling of germin reported by Gane *et al.* (1998).

Phosphomannose isomerases

Phosphomannose isomerases (PMIs) (EC 5.3.1.8) are enzymes that catalyse the interconversion of mannose-6-phosphate and fructose-6-phosphate. Although they are considered to be zinc-containing metalloproteins, a Fe(III)-hydroxyphenylalanine site has been identified in the PMI from *Candida albicans* when expressed in *E. coli* (Proudfoot *et al.*, 1996; Smith *et al.*, 1997). On the basis of sequence comparison, the PMIs have been divided into three classes (Proudfoot *et al.*, 1994), within which the class II enzymes are involved in a variety of pathways including capsular polysaccharide biosynthesis and D-mannose metabolism. They all include, toward their C-terminus, two histidine-containing motifs similar in sequence and spacing to those found in the other proteins discussed in this review. One difference is that the histidines of the first motif are often displaced to give a $G(X)_7HXH(X)_9G$ sequence.

Notable amongst these enzymes, many of which are found in pathogenically important organisms, are the spore polysaccharide protein (P39631) from *Bacillus subtilis*, the Cap5F and Cap8F proteins from *Staphylococcus aureus* (Sau and Lee, 1996) and the closely related (>90%) but possibly misnamed 'nucleotide binding' protein (S33518) from the mycoplasma *Acholeplasma laidlawii*. Of this large group of related proteins, the smallest (128 AA) is the *Synechocystis* PMI gene (gil1001180). Others include: a sequence (S04426) from *Synechococcus*; a sequence (created as SwissProt virtual sequence VIRT2200) in the upstream region of the aspartate racemase gene (D84067) from the archaeon *Desulfurococcus* (Yohda *et al.*, 1996) (see section on cryptic genes below); GDP-mannose pyrophosphorylase (Q46859) from *Yersinia enterocolitica;* and P07874 – the 56KD bifunctional enzyme from *Pseudomonas aeruginea* with both PMI and GDP-mannose pyrophosphorylase activities (May *et al.*, 1994). This latter enzyme is involved in the polymerisation of the viscous mucoid exopolysaccharide, alginate – a compound composed of 1,4-linked α-L-guluronic acid and β-D-mannuronic acid subunits. Its pathogenic significance lies in the fact that it protects the bacterial cell from host immune responses and antibiotics, and thus is also the major cause of mortality in patients suffering from cystic fibrosis. In contrast to their role in pathogenesis, bacterially produced alginates, along with

those from marine microalgae (seaweeds), are of great economic importance in the food industry (Renn, 1997).

There are also eukaryotic equivalents of these enzymes (e.g S53039 from yeast) in which either one or both motifs are conserved. Related sequences from *Caenorhabditis elegans* include two with unknown function, gil1707132, and gil1082118. It had been noted previously (see attachment to the sequence submission) that the latter sequence contained a domain with weak similarity to the acidic domain of a rice glutelin precursor. In fact, it is more similar to a region of the legumin precursor (S54207) from *Magnolia salicifolia* and other species.

Polyketide synthases

Amongst the bacterial sequences with the closest homology to germin is *curC* (Q02586) from *Streptomyces cyaneus* (Bergh and Uhlén 1992a,b). The gene product is probably a cyclase (although its exact function is unknown) and is part of the synthetic pathway of the antibiotic curamycin. This compound has a polyketide skeleton consisting of orsellinic acid – the simplest aromatic polyketide possible. Closely related genes with similar function from other *Streptomyces* species include those from *S. coelicolor* (X55942) (Davis and Chater, 1990), *S. halstedii* (L05390) (Blanco *et al.*, 1993), and *S. glaucescens* (M80674) (Bibb *et al.*, 1989) and X77865 (see *Figure 4* in Blanco *et al.*, 1993). This latter study also identified sequences closely related to PGERISEHYHPYSE (residues 47–60 in L05390) (cf. motif 1) that are conserved amongst several β-glucosidases from plants, fungi, and bacteria (e.g. P26208 from *Clostridium thermocellum*).

Additional members of this group of putative cyclases include the *B. subtilis* sequence P54430 and the sequence *orf1* (gil1402862) (product Pep1) within U60777 (Smith *et al.* 1993) from the cryptic transposon Tn4321, a part of the broad host range IncP beta plasmid R751 of *Enterobacter aerogenes*. This Pep1 sequence (nucleotides 1704–2126) is situated between a transposase *tnpA*, and *orf2* (product Pep2), a putative member of the LysR family of bacterial regulator proteins. The notes accompanying the submission, whilst noting 'possible polyketide cyclase on basis of weak similarity to *TcmJ* of *Streptomyces glaucescens*, M80674', did not recognize that the closest protein relative to Pep1, as revealed by a BLASTP search, is an Arabidopsis germin (U75187).

Another possible member of this class is the 143 AA protein encoded by sequence gil1653078 from *Synechocystis*. The smallest (110–132 AA) proteins in this group are two from *Erwinia chrysanthemi*, first a 'pectin degradation protein KDGF' (accession Q05527) (Condemine and Robert-Baudouy, 1991) and secondly the sequence gil1772621 in L39897, one from *Mycobacterium tuberculosis* (gil1654019) (Philipp *et al.*, 1996), and one from *Bacillus subtilis* (gil1881251).

Epimerases

Sequence analysis has revealed that the cupin protein superfamily include a series of epimerases involved in the synthesis of bacterial exopolysaccharides. These enzymes include dTDP-4-dehydrorhamnose 3,5-epimerases such as the ExpA8 protein recently shown to be involved in the biosynthesis of galactoglucan (exopolysaccharide

II) in *Rhizobium melioti* (Becker *et al.*, 1997). The complete list of these enzymes includes examples from *E. coli* (Stevenson *et al.*, 1994; Yao and Valvano, 1994; Marolda and Valvano, 1995), *Xanthomonas campestris* (Koplin *et al.*, 1993) *Leptospira interrogans* (Mitchison *et al.*, 1997), *Neisseria gonorrhoeae* (Robertson *et al.*, 1994; Petering *et al.*, 1996), *N. meningitidis* (Hammerschmidt *et al.*, 1994), *Yersinia pseudotuberculosis* (Thorson *et al.*, 1994), *Sphingomonas* (Yamazaki *et al.*, 1996), *Streptomyces griseus* (Pissowotzki *et al.*, 1991; Krugel *et al.*, 1993), *Salmonella typhimurium* (Jiang *et al.*, 1991; Wang *et al.*, 1992; Zhang *et al.*, 1993), *Shigella flexneri* (Macpherson *et al.*, 1994), *Haemophilus actinomycetemcomitans* (gil1944162) and *Synechocystis* (gil165678). The equivalent gene from *C. elegans* is gil1049348.

A similar enzyme dTDP-6-deoxy-L-mannose-dehydrogenase (gil1651977) is known from *Synechocystis*.

Dioxygenases

Amongst these iron-containing enzymes is one class which contains the two histidine-containing motifs, although the first motif has only a single histidine residue. This specific class of enzyme comprise the 3-hydroxyanthranilate 3,4-dioxygenases (3-HAO) that catalyse the synthesis of quinolinic acid from 3-hydroxyanthranilic acid. Examples include a yeast sequence (accession P47096) and similar proteins from *C. elegans* (Z70755) and rat (D44494) (Malherbe *et al.*, 1994). Additionally, the cysteine dioxygenases such as that from rat (P21816) (Hosokawa *et al.*, 1990) clearly contain an equivalent first motif, though the second is either absent, or present as a weakly similar motif separated by a gap of approximately 40 residues.

Related proteins with no known function

Additional analysis identified a series of progressively smaller proteins, each of which contain motifs 1 and 2 (Dunwell and Gane, 1998). Included amongst these proteins, none of which has any known function, are two similar accessions (D90909, D90910) from *Synechocystis*, the first of which has the Arabidopsis germin X91957 as the closest relative, and the second of which is similar to cytochrome c551 from *Rhodococcus*. Slightly shorter sequences include one from *Mycobacterium tuberculosis* (gil2104394), two from the methanogenic archaeon *Methanococcus jannaschii* (Bult *et al.*, 1996), [gil1592216 (U67602) from predicted coding region MJ1618 and its close relative gil1499583 (U67521)], and two *E. coli* sequences, P38522 which is similar to an aldehyde dehydrogenase (Heim and Strehler, 1991), and its longer version – the immunity repressor protein (D90768).

The two sequences from *M. jannaschii* have not yet been assigned a function, though it is to be expected that the type of metabolic reconstruction being devised by Selkov *et al.* (1997) will help towards that end.

The smallest (77 AA) of all those identified is a membrane spanning protein (MSP) (Li and Strohl, 1996) (U37580) from *Streptomyces coelicolor*. This protein, which has similarity to *E. coli* PMI, and also to an upstream sequence (SwissProt virtual sequence VIRT17611) in the prismane gene (Z11975) from *Desulfovibrio desulfuricans* (Stokkermans *et al.*, 1992) (see next section), is the putative progenitor of all these two motif proteins, and subsequently of all the related cupin superfamily of proteins.

Hypothetical cryptic genes

There are several examples of previously unknown, germin-like genes revealed recently (J.M. Dunwell, unpublished results) by means of TBLASTN searches, which are able to reveal putative coding regions in the upstream regions of other genes. The most likely explanation for these cryptic ORFs, which may be found on either strand in any reading frame, is the accidental ligation of DNA sequences during the cloning procedures employed in the construction of cDNA libraries (Bhatia *et al.*, 1997).

Of the six examples described here, the most complete is that found in the upstream sequence of the *Arabidopsis* polyubiquitin gene *ubq4* (Burke *et al.*, 1988). The 4024 nucleotide cDNA sequence (gi|987518) from this study shows the coding region to start at position 2649. Inspection of the 'upstream' sequence reveals that nucleotides 1195 to 1873, with the insertion of a 'g' at position 1288 in order to overcome a frame shift error introduced by the sequencing procedure, encodes a germin-like protein identical to gi|1755168 and that encoded by the Arabidopsis EST, ATTS 0248 (gi|16847).

A very similar situation is present in the upstream sequence of a groESL operon (D12677) from *Synechocystis* (Lehel *et al.*, 1993). Nucleotides 3–208, with addition of an additional 'g' at position 104, encode a protein identical (with the exception of a single Asp to Tyr mutation) to the C-terminal 69 amino acids encoded by gi|1652486 (accession D90905), a PMI (see above) similar to gi|1001180 from the same species. Perhaps coincidentally, it is reported that the correct folding of a recombinant type I PMI from *Candida albicans* (Cleasby *et al.*, 1996) in *E. coli* is increased by the co-expression of the GroES or GroEL chaperones (Proudfoot *et al.*, 1996).

The third example (VIRT28811) (*Table 4*) is encoded by nucleotides 338–1 (negative strand) of the upstream sequence of gi|560029 which contains at nucleotide position 507–2363 a gene for a xylan 1,4-beta-xylosidase (Oh and Choi, 1994). Intriguingly, the closest neighbour of this cryptic sequence is the next sequence described.

The fourth example (VIRT2200) (*Table 4*) is found in the positive strand (nucleotides 104–450) of the upstream sequence (although not in the same reading frame) of the aspartate racemase gene from the archaeon *Desulfurococcus* strain SY (Yohda *et al.*, 1996). This sequence is most similar (50% similar over a 79 AA region) to the *rfb* gene from *Salmonella enterica* (see section on epimerases above).

The fifth example (VIRT17611) (*Table 4*) is that from nucleotides 238–727 in the upstream sequence (same reading frame) of the prismane protein from *Desulfovibrio vulgaris* (Stokkermans *et al.*, 1992). Again, this derived sequence has another cryptic sequence as its closest neighbour, in this instance a protein (VIRT20248) (*Table 4*) from *Pseudomonas lemoignei* encoded by the nucleotides 1618–1286 (negative strand) of accession U12977, between a glycerol-3-phosphate-dehydrogenase (gi|531467) and a poly(3-hydroxybutyrate) depolymerase A (gi|531466) (Jendrossek *et al.*, 1995).

TWO-DOMAIN PROTEINS (BI-CUPINS)

Oxalate decarboxylase and related microbial proteins

The principal organisms responsible for the biodegradation of wood are white-rot and brown-rot basidiomycete fungi (Dutton *et al.*, 1993; Dutton and Evans, 1996).

Table 4. Putative cupin sequences encoded by cryptic genes in upstream 'non-coding' regions of other bacterial genes (J.M. Dunwell, unpublished results)

>VIRT28811, 112 AA (*Bacillus stearothermophilus*)
FTFKLFTTAFAGRLRRSGRIRRHSHPLFEVNLLLAGSQVMTVGGKRYVQQPGDLLLVRPEDVHESRALGTEPMT
YYCLHFNVDEPVFASYCAVARTRSTPRTGRWPRSFGQA

>VIRT2200, 115 AA (*Desulfurococcus*)
WGERVKAEIKNLIDRGTYRKLPLFEGELPEGSYAQIVEVKPGQTVKKHYHLHQYELFYIMSGEARLGIGETGYL
ARPGDIFLVKPKTVHWVINEKEEPFRLFVVKLNYKGDDSVW

>VIRT17611 163 AA (*Desulfibrio desulfuricans*)
RPAAHQRKKHLPASSRPPERILRLFCSVTIVTDVHTPPWYTGIIKSKEGTMQRLLKNLDYATALPLAAQVDCQP
GQVASKTLVQNDAVGITLFAFDKGEKISAHTSTGDAFVLALEGQGQVTINGQTSPLKAGESIIMPAGQPHSVSA
VERFKMLLVVIFPPE

>VIRT20248 111 AA (*Pseudomonas lemoignei*)
MALPHASSGQLIDVRPLGSQLTSAPSRAILKTSGLELMHMVLPAGRTVPEHRVPGECTIQCIEGSVELTSHGNT
QLMRAGDLVYFEGGVVHALHAQEDSSLLVTILMKHEA

Whereas the former type are able to degrade all components of the wood cell wall, brown-rot fungi remove the cellulosic fraction and leave the lignin undegraded. It has been known for many years that most brown-rot fungi accumulate large quantities of oxalic acid in certain culture media (Shimazono, 1955). In similar conditions, the white rot organisms fail to show this accumulation, because of the activity of oxalate decarboxylase (ODC)-the enzyme (EC 4.1.1.2) that converts oxalic acid to formate and carbon dioxide. (It has recently been shown by Micales (1995) that brown-rot fungi also produce an ODC enzyme.)

In commercial terms, this enzyme has potential value (Hiatt and Owades, 1987) as an additive that can be supplied during the brewing process in order to reduce the level of oxalate (to prevent 'gushing') (Haas and Fleischman, 1961) and to degrade 'beerstone', crystals of calcium oxalate formed in the vat and therefore likely to obstruct pipe-work.

The *Collybia velutipes* ODC enzyme was first characterized in detail and cloned by Mehta and Datta (1991) and its sequence recently published (Datta *et al.*, 1996). Analysis of the protein sequence (Dunwell and Gane, 1998) has now revealed the surprising fact that the protein is a duplicated one, with the two domains each showing sequence similarity to the ancestral germin protein (Bäumlein *et al.*, 1995). It is therefore quite closely related to the plant storage proteins such as vicilin and legumin, and is also related to the other oxalate degrading enzyme ox–ox (see above). Use of the germin homology model (Gane *et al.*, 1998) should enable predictions to be made of the structure of the active site of this ODC (Labrou and Clonis, 1995) in the near future.

Identification of this two-domain ODC in the fungus *C. velutipes* has also led to the discovery of a closely related bacterial sequence, namely gi|1652630 from the unicellular cyanobacterium *Synechocystis* PCC6303 (Kaneko *et al.*, 1996). This hypothetical protein is encoded by nucleotides 15268–16452 of accession D90907 and its closest neighbour as estimated by residues 250–381 is Arabidopsis germin-like protein U95046 (30% identical, 49% similar). Another example of duplication is provided by a protein (accession P42106) (Yoshida *et al.*, 1995) of unknown function from *Bacillus subtilis*.

Despite the fact that the critical residues within each motif, and the spacing between motifs, have been conserved in all three examples, there are a number of significant differences between composition of these proteins. First, the overlaps of the duplicated regions are of different length. Additionally, BLASTP searches with BEAUTY annotation (Worley *et al.*, 1990) show that the stretches of sequence showing significant homology with other proteins are located primarily in the C-terminal section of oxalate decarboxylase, whereas they are located mostly in the N-terminal half of accession P42106. The third protein, gil1652630, has equal similarity (particularly to germins/vicilins) in each half.

In terms of their evolution, there are a number of possible origins for this duplication, which previously has never been found outside the plant kingdom. The simplest hypothesis is that the two domains of each protein have diverged differentially after a single duplication event. Alternatively, duplication might have occurred three times, most recently in *Synechocystis* (gil1652630) in which the two domains show equal similarity to other proteins. The third possibility is that this latter protein is the only one to be a product of duplication, and that the other two examples have instead evolved from a homologous recombination event between two similar DNA sequences. In this case, gil1652630 would be the direct and sole ancestor of the higher plant storage proteins (32/52 of its nearest neighbours are such proteins).

Sucrose binding protein (SBP)

The fact that sucrose-binding proteins (or at least the example from soybean, Q04672 – Grimes *et al.*, 1992) is a two-domain member of the germin family, was first reported by Braun *et al.* (1996) in a study of a vicilin-like protein (Z50791) from the cycad *Zamia furfuracea*. In a detailed analysis of sequence, these authors suggest that the replacement of a conserved proline residue (found in vicilins) with an isoleucine in the SBP C-terminal domain might create stearic hindrances and therefore prevent the trimerization of subunits typical of storage proteins.

Storage proteins

Seed storage globulins are divided into two types, the vicilin-like 7S, and the legumin-like 11S. The former 7S proteins are trimers assembled from two main types of 50 or 70 kD subunits, whereas the subunits of 11S proteins are synthesized as a precursor which is processed into an acidic α- and a basic β-chain whilst remaining connected by a disulphide bond. These subunits are then assembled to produce a hexameric holoprotein. Within the last few years it has been found that the genes coding for both the legumin and vicilin subunits are the result of an internal duplication of a single domain ancestral gene (Lawrence *et al.*, 1994). It was shown subsequently (Bäumlein *et al.*, 1995) that the spherulins and wheat germins might also have evolved from this common 'protogermin' ancestor, for which various bacterial candidates have been proposed by Dunwell and Gane (1998).

Role of glycosylation

GERMIN AND GERMIN-LIKE PROTEINS

All examples of germin and its close relatives are glycosylated. It is notable, however, that the glycosylation site(s) vary in position within the protein, and perhaps could be used as a diagnostic feature to categorise the different classes of these proteins. The only detailed study of the glycan side chains was that conducted on the wheat germins (Jaikaran *et al.*, 1990).

Several of the other single-domain germin-like proteins, such as the ABPs, also contain glycosylation sites, with the position of such sites being a feature of each group of protein.

STORAGE PROTEINS

Phaseolin, the major storage protein of bean, is encoded by a small gene family divided into two highly homologous classes, α and β (Slightom *et al.*, 1985). Polypeptides of both classes are synthesized as two glycoforms bearing either one or two glycan side chains. This glycosylation is not required for the assembly of the monomers into the final trimeric structure (Vitale *et al.*, 1995) but it does affect the rate of assembly. More recently, it has been shown that glucose trimming of the side chains is a key event in the assembly process (Lupatelli *et al.*, 1997).

Structure predictions

As described above, germins possess sequence similarity with legumin and vicilin storage proteins. Recently, this information was utilized to generate a 3D model of germin (Gane *et al.*, 1998; see *Figure 1*) from which a potential ox–ox active site could be predicted. This site, comprising a cluster of three histidine residues, was located within the conserved β-barrel structure and was similar to those found in copper amine oxidases, although it should be noted that the metal cofactor of ox–ox has never been unequivocally demonstrated. In addition, the modelling study used the potential for charge-charge interactions between monomers to produce a pentameric assembly, in which the β-barrels remain parallel to the five-fold axis, thereby maintaining solvent exposure for the predicted active sites.

This pentameric structure of germin is reminiscent of that found in the pentraxins, a family of vertebrate plasma proteins which include human C-reactive protein, the classical cytokine-regulated acute phase protein produced in response to inflammation, and universally measured as an objective marker of disease activity (Emsley *et al.*, 1994; Gewurz *et al.*, 1995; Hohenester *et al.*, 1997; Srinavasan *et al.*, 1994). Like germin, the pentraxins are protease resistant, and structural analysis reveals another similarity in that their tertiary fold is very similar to that of another class of plant storage protein, namely the legume lectins. Recently, several groups (Goodman *et al.*, 1996; Basile *et al.*, 1997) have identified an additional class of 'long pentraxins'. Although only the C-terminus shows features of the pentraxin family, they are still thought to form pentameric complexes. In some ways these extended proteins are analogous to the two-domain, bi-cupin proteins.

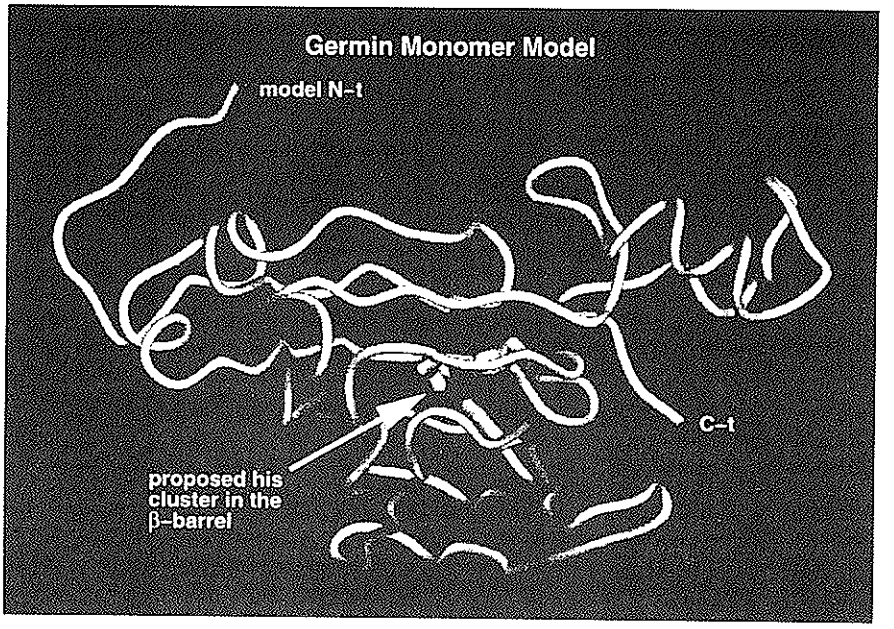

Figure 1. Proposed molecular model of the germin molecule, based on the homology with the storage protein vicilin (Gane *et al.*, 1998). The arrow indicates the position of the histidine cluster, the suggested active site of the oxalate oxidase (courtesy Jim Warwicker).

It is expected that the 3D model of germin will soon be used to generate structural predictions of other proteins in the superfamily.

Concluding remarks

The cereal germin proteins, along with their close relatives from gymnosperms and dicotyledonous plants, have emerged recently from a position of being considered as interesting proteins with no known function, to being identified as enzymes concerned with cell wall structure (Lane, 1994), fungal defence (Dumas *et al.*, 1993; Zhang *et al.*, 1995; Hurkman and Tanaka, 1996b), salt tolerance (Michalowski and Bohnert, 1992; Hurkman and Tanaka, 1996a), floral induction (Heintzen *et al.*, 1994; Ono *et al.*, 1996), and as markers of somatic embryogenesis (Domon *et al.*, 1995). In addition, knowledge of their ancestral relationship to plant storage proteins (Dunwell and Gane, 1998) has been used to generate a 3D model (Gane *et al.*, 1998) and identify details of the potential active site. Together with the identification of several other related proteins in different enzyme classes – a finding predicted by Lane (1994) and similar to the results of Murzin (1993) and Babbit *et al.* (1995) – these new observations will undoubtedly lead to a further rapid expansion in fundamental and applied interest in this most fascinating group of proteins.

Acknowledgements

I would like to thank the Biotechnology and Biological Sciences Research Council,

and ZENECA plc for funding. I am also most grateful to Dr Jim Warwicker for provision of *Figure 1*.

References

ADAMS, P., THOMAS, J.C., VERNON, D.M., BOHNERT, H.J. AND JENSEN, R.G. (1992). Distinct cellular and organismic responses to salt stress. *Plant and Cell Physiology* **33**, 1215–1223.

ALBRECHT, S., BRANDL, H. AND SCHÖNFELS, C. (1994). Human oxalate – really just an end-product of metabolism? *Angewandte Chemie International Edition English* **33**, 1780–1781.

ANDOLFATTO, P., BORNHAUSER, A., BOHNERT, H.J. AND THOMAS, J.C. (1994). Transformed hairy roots of *Mesembryanthemum crystallinum*: gene expression patterns upon salt stress. *Physiological Plant Physiology* **90**, 708–714.

ANDREWS, J.C. AND VISER, E.T. (1951). The oxalic acid content of some common foods. *Food Research* **16**, 306–312.

ANONYMOUS (1992). These rhubarb-eating bugs like sodium oxalate too. *Chemical Engineering*, October, p. 17.

BABBITT, P.C., MRACHKO, G.T., HASSON, M.S., HUISMAN, G.W., KOLTER, R., RINGE, D., PETSKO, G.A., KENYON, G.L. AND GERLT, J.A. (1995). A functionally diverse enzyme superfamily that abstracts the α protons of carboxylic acids. *Science* **267**, 1159–1161.

BAETZ, A.L. AND ALLISON, M.J. (1989). Purification and characterisation of oxalyl-coenzyme A decarboxylase from *Oxalobacter formigenes*. *Journal of Bacteriology* **171**, 2605–2608.

BAILEY, T.L. AND ELKAN, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the 2nd International Conference on Intelligent Systems for Molecular Biology*, AAAI Press, Menlo Park, California. pp. 28–36.

BALION, C.M. AND THIBERT, R.J. (1994). Determination of oxalate by luminol chemiluminescence. *Clinical Chemistry* **40**, Abstr. 0573.

BASILE, A., SICA, A., D'ANIELLO, E., BREVIARIO, F., GARRIDO, G., CASTELLANO, M., MANTOVANI, A. AND INTRONA, M. (1997). Characterization of the promoter for the human long pentraxin PTX3. Role of NF-κB in tumor necrosis factor-α and interleukin-1β regulation. *Journal of Biological Chemistry* **272**, 8172–8178.

BATICH, C. AND VAGHEFI, F. (1994). Process for microencapsulating cells. US Patent 5286495.

BÄUMLEIN, H., BRAUN, H., KAKHOVSKAYA, I.A. AND SHUTOV, A.D. (1995). Seed storage proteins of spermatophytes share a common ancestor with desiccation proteins of fungi. *Journal of Molecular Evolution* **41**, 1070–1075.

BECKER, A., RUBERG, S., KUSTER, H., ROXLAU, A.A. AND KELLER, M. (1997). The 32-kilobase *exp* gene cluster of *Rhizobium melioti* directing the biosynthesis of galactoglucan: genetic organisation and properties of the encoded gene products. *Journal of Bacteriology* **179**, 1375–1384.

BERGH, S. AND UHLÉN, M. (1992a). Analysis of a polyketide synthesis-encoding gene cluster of *Streptomyces curacoi*. *Gene* **117**, 131–136.

BERGH, S. AND UHLÉN, M. (1992b). Cloning, analysis, and heterologous expression of a polyketide synthase gene cluster of *Streptomyces curacoi*. *Biotechnology and Applied Biochemistry* **15**, 80–89.

BERNA, A. AND BERNIER, F. (1997). Regulated expression of a wheat germin gene in tobacco: oxalate oxidase activity and apoplastic localization of the heterologous protein. *Plant Molecular Biology* **33**, 417–429.

BERNIER, F., LEMIEUX, G. AND PALLOTTA, D. (1987). Gene families encode the major encystment-specific proteins of *Physarum polycephalum* plasmodia. *Gene* **59**, 265–277.

BHATIA, U., ROBISON, K. AND GILBERT, W. (1997). Dealing with database explosion: a cautionary note. *Science* **276**, 1724–1725.

BIBB, M.J., BIRO, S., MOTAMEDI, H., COLLINS, J.F. AND HUTCHINSON, C.R. (1989). Analysis of the nucleotide sequence of the *Streptomyces glaucescens tcm1* genes provides key information about the enzymology of polyketide antibiotic synthesis. *EMBO Journal* **8**, 2727–2736.

BLANCO, G., BRIAN, P., PEREDA, A., MÉNDEZ, C., SALAS, J.A. AND CHATER, K.F. (1993). Hybridization and DNA sequence analyses suggest an early evolutionary divergence of related biosynthetic gene sets encoding polyketide antibiotics and spore pigments in *Streptomyces* spp. *Gene* **130**, 107–116.

BOHNERT, H.J., OSTREM, J.A., CUSHMAN, J.C., MICHALOWSKI, C.B., RICKERS, J., MEYER, G., DEROCHER, E.J., VERNON, D.M., KRUEGER, M., VASQUEZ-MARTIN, L., VELTEN, J., HOEFNER, R. AND SCHMITT, J.M. (1988). *Mesembryanthemum crystallinum*, a higher plant model for the study of environmentally induced changes in gene expression. *Plant Molecular Biology Reporter* **6**, 10–28.

BRAUN, H., CZIHAL, A., SHOTOV, A.D. AND BAUMLEIN, H. (1996). A vicilin-like seed protein of cycads: similarity to sucrose-binding proteins. *Plant Molecular Biology* **31**, 35–44.

BROWN, J.C. AND JONES, A.M. (1994). Mapping the auxin-binding site of auxin-binding protein 1. *Journal of Biological Chemistry* **269**, 21136–21140.

BULT C.J., WHITE, O., OLSEN, G.J., ZHOU, L., FLEISCHMANN, R.D., SUTTON, G.G., BLAKE, J.A., FITZGERALD, L.M., CLAYTON, R.A., GOCAYNE, J.D., KERLAVAGE, A.R., DOUGHERTY, B.A., TOMB, J-F., ADAMS, M.D., REICH, C.I., OVERBEEK, R., KIRKNESS, E.F., WEINSTOCK, K.G., MERRICK, J.M., GLODEK, A., SCOTT, J.L., GEOGHAGEN, N.S.M., WEIDMAN, J.F., FUHRMAN, J.L. PRESLEY, E.A., NGUYEN, D., UTTERBECK, T.R., KELLEY, J.M., PETERSON, J.D., SADOW, P.W., HANNA, M.C., COTTON, M.D., HURST, M.A., ROBERTS, K.M., KAINE, B.P., BORODOVSKY, M., KLENK, H-P., FRASER, C.M., SMITH, H.O., WOESE, C.R. AND VENTER, J.C. (1996). Complete genome sequence of the methanogenic achaeon *Methanococcus jannaschii*. *Science* **273**, 1058–1073.

BURKE, T.J., CALLIS, J. AND VIERSTRA, R.D. (1988). Characterisation of a polyubiquitin gene from *Arabidopsis thaliana*. *Molecular and General Genetics* **213**, 435–443.

BUSTOS, S.A. AND SCHLEIF, R.F. (1993). Functional domains of the AraC protein. *Proceedings of the National Academy of Sciences USA* **90**, 5638–5642.

CHANDRA, T.S. AND SHETHNA, Y.I. (1975). Oxalate and formate in *Alcaligenes* and *Pseudomonas* species. *Antonie Van Leewenhoek* **41**, 465–477.

CHIRIBOGA, J. (1966). Purification and properties of oxalic acid oxidase. *Archive of Biochemistry and Biophysics* **116**, 516–523.

CHOI, S.Y. (1996). Molecular cloning and expression of the hot pepper *ERabp1* gene encoding auxin-binding protein. *Plant Molecular Biology* **32**, 995–997.

CLEASBY, A., WONACOTT, A., SKARZYNSKI, T., HUBBARD, R.E., DAVIES, G.J., PROUDFOOT, A.E., BERNARD, A.R., PAYTON, M.A. AND WELLS, T.N. (1996). The X-ray crystal structure of phosphomannose isomerase from *Candida albicans* at 1.7 angstrom resolution. *Nature Structural Biology* **3**, 470–479.

CONDEMINE, G. AND ROBERT-BAUDOUY, J. (1991). Analysis of an *Erwinia chrysanthemi* gene cluster involved in pectin degradation. *Molecular Microbiology* **5**, 1191–2202.

CORNICK, N.A. AND ALLISON, M.J. (1996). Assimilation of oxalate, acetate, and CO_2 by *Oxalobacter formigenes*. *Canadian Journal of Microbiology* **42**, 1081–1086.

DATTA, A., MEHTA, A. AND NATARAJAN, K. (1996). Oxalate decarboxylase. US Patent 5547870.

DAVIS, N.K. AND CHATER, K.F. (1993). Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. *Molecular Microbiology* **4**, 1679–1691.

DATTA, P.K. AND MEEUSE, B.J. (1955). Moss oxalic acid oxidase-a flavo-protein. *Biochimica Biophysica Acta* **17**, 602–603.

DELSENY, M., GAUBIER, P., HULL, G., SAEZ-VASQUEZ, J., GALLOIS, P., RAYNAL, M., COOKE, R. AND GRELLET, F. (1994). Nuclear genes expressed during seed desiccation: relationship with responses to stress. In *Stress-Induced Gene Expression in Plants*, ed. A.S. Basra, pp. 25–59. Switzerland: Harwood Academic Publishers.

DE PAUW, L. AND TOUSSAINT, C. (1996). Primary hyperoxaluria. *Revue Medicale de Bruxelles* **17**, 67–74.

DICKMAN, M.B. AND MITRA, A. (1992). *Arabidopsis thaliana* as a model for studying *Sclerotinia sclerotiorum* pathogenesis. *Physiological and Molecular Plant Pathology* **41**, 255–263.

DOMON, J-M., DUMAS, B., LAINÉ, E., MEYER, Y., DAVID, A. AND DAVID, H. (1995). Three glycosylated polypeptides secreted by several embryogenic cell lines of pine show highly specific serological affinity to antibodies directed against the wheat germin apoprotein monomer. *Plant Physiology* **108**, 141–148.

DRATEWKA-KOS, E., RAHMAN, S., GRZELCZAK, Z.F. AND KENNEDY, T.D. (1989). Polypeptide structure of germin as deduced from cDNA sequencing. *Journal of Biological Chemistry* **264**, 4896–4900.

DUBCOVSKY, J., LUO, M-C. AND DVORÁK, J. (1995). Linkage relationships among stress-induced genes in wheat. *Theoretical and Applied Genetics* **91**, 795–801.

DUMAS, B., FREYNISSET, G. AND PALLET, K. (1995). Tissue-specific expression of germin-like oxalate oxidase during development and fungal infection of barley seedlings. *Plant Physiology* **107**, 1091–1096.

DUMAS, B., SAILLAND, A., CHEVIET, J-P., FREYSSINET, G. AND PALLETT, K. (1993). Identification of barley oxalate oxidase as a germin-like protein. *Comptes Rendus de l'Academie des Sciences, Paris* **316**, 793–798.

DUNWELL, J.M. AND GANE, P.J. (1998). Microbial ancestors of seed storage proteins:conservation of motifs in a functionally diverse superfamily of enzymes. *Journal of Molecular Evolution* (in press).

DUTTON, M.V. AND EVANS, C.S. (1996). Oxalate production by fungi: its role in pathogenicity and ecology in the soil environment. *Canadian Journal of Microbiology* **42**, 881–895.

DUTTON, M.V., EVANS, C.S., ATKEY, P.T. AND WOOD, D.A. (1993). Oxalate production by Basidiomycetes, including the white-rot species *Coriolus versicolor* and *Phanerochaete chrysosporium*. *Applied Microbiology and Biotechnology* **39**, 5–10.

EDGERTON, M.D., TROPSHA, A. AND JONES, A.M. (1994). Modelling the auxin-binding site of auxin-binding protein 1 of maize. *Phytochemistry* **35**, 1111–1123.

EMSLEY, J., WHITE, H.E., O'HARA, B.P., OLIVA G., SRINAVASAN, N., TICKLE, I.J., BLUNDELL, T.L., PEPYS, M.B. AND WOOD, S.P. (1994). Structure of pentameric human serum amyloid P component. *Nature* **367**, 338–345.

FARQUHARSON, G.J., KILDEA, J.D., GROSS, A.E. AND GROCOTT, S.C. (1995). Liquor oxalate stabilizers. US Patent 5385586.

FELLNER, M., EPHRITIKHINE, G., BARBIER-BRYGOO, H. AND GUERN, J. (1996). An antibody raised to a maize auxin-binding protein has inhibitory effects on cell division of tobacco protoplasts. *Plant Physiology and Biochemistry* **34**, 133–138.

FINLAYSON, B. AND PECK, A. (1988). Gene therapy for metabolic disorders. PCT Patent Application WO 88/08450.

FRANKE, W. AND HASSE, K. (1937). Zur biologische Oxydation der Oxalsäure. *Zeitschrift für Physiologische Chemie* **249**, 231–255.

GANE, P.J., DUNWELL, J.M. AND WARWICKER, J. (1997). Modelling based on the structure of vicilins predicts a histidine cluster in the active site of oxalate oxidase. *Journal of Molecular Evolution* (in press).

GEWURZ, H., ZHANG, X.H. AND LINT, T.F. (1995). Structure and function of the pentraxins. *Current Opinions in Immunology* **7**, 54–64.

GODOY, G., STEADMAN, J.R., DICKMAN, M.B. AND DAM, R. (1990). Use of mutants to demonstrate the role of oxalic acid in pathogenicity of *Sclerotinia sclerotiorum* on *Phaseolus vulgaris*. *Physiological and Molecular Pathology* **37**, 179–191.

GOODMAN, A.R., CARDOZO, T., ABAGYAN, R., ALTMEYER, A., WISNIEWSKI, H.G. AND VILCEK, J. (1996). Long pentraxins:an emerging group of proteins with diverse functions. *Cytokine and Growth Factor Review* **7**, 191–202.

GRIMES, H.D., OVERVOORDE, P.J., RIPP, K., FRANCESCHI, V.R. AND HITZ, W.D. (1992). A 62-kD sucrose binding protein is expressed and localized in tissues actively engaged in sucrose transport. *The Plant Cell* **4**, 1561–1574.

GRZELCZAK, Z.F. AND LANE, B.G. (1984). Signal resistance of a soluble protein to enzyme hydrolysis. An unorthodox approach to the isolation and purification of germin, a rare growth-related protein. *Canadian Journal of Biochemistry and Cell Biology* **62**, 1351–1353.

GRZELCZAK, Z.F., RAHMAN, S., KENNEDY, T.D. AND LANE, B.G. (1985). Germin.

Compartmentation of the protein, its translatable mRNA and its biosynthesis among roots, stems, and leaves of wheat seedlings. *Canadian Journal of Biochemistry and Cell Biology* **63**, 1003–1013.

HAAS, G.J. AND FLEISCHMAN, A.I. (1961). The rapid enzymatic determination of oxalate in wort and beer. *Journal of Food and Agriculture* **9**, 451–452.

HAMMERSCHMIDT, S., BIRKHOLZ, C., ZAHRINGER, U., ROBERTSON, B.D., VAN PUTTEN, J., EBELING, O. AND FROSCH, M. (1994). Contribution of genes from the capsule gene complex (cps) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. *Molecular Microbiology* **11**, 885–896.

HANSEN, E.H., WINTHER, S.K. AND GUNDSTRUP, M. (1994). Enzymatic assay of oxalate in urine by flow injection analysis using immobilized oxalate oxidase and chemiluminescence detection. *Analytical Letters* **27**, 1239–1253.

HEIM, R. AND STREHLER, E.E. (1991). Cloning an *Escherichia coli* gene encoding a protein remarkably similar to mammalian aldehyde dehydrogenases. *Gene* **99**, 15–23.

HEINTZEN, C., FISCHER, R., MELZER, S., KAPPELER, S., APEL, K. AND STAIGER, D. (1994). Circadian oscillations of a transcript encoding a germin-like protein that is associated with cell walls in young leaves of the long-day plant *Sinapis alba* L. *Plant Physiology* **106**, 905–915.

HENIKOFF, S. AND HENIKOFF, J.G. (1991). Automated assembly of protein blocks for database searching. *Nucleic Acids Research* **19**, 6565–6572.

HIATT, W.R. AND OWADES, J.L. (1987). Oxalic acid removal in beer production. US Patent 4652452.

HÖFTE, H., DESPREZ, T., AMSELEM, J., CHIAPPELLO, H., CABOCHE, M., MOISAN, A., JOURJON, M-F., CHARPENTEAU, L., BERTHOMIEU, P., GUERRIER, D., GIRAUDAT, J., QUIGLEY, F., THOMAS, F., YU, D-Y., MACHE, R., RAYNAL, M., COOKE, R., GRELLET, F., DELSENY, M., PARMENTIER, Y., DE MARCILLAC, G., GIGOT, C., FLECK, J., PHILIPPS, G., AXELOS, M., BARDET, C., TREMOUSAYGUE, D. AND LESCURE, B. (1993). An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. *The Plant Journal* **4**, 1051–1061.

HOHENESTER, E., HUTCHINSON, W.L., PEPYS, M.B. AND WOOD, S.P. (1997). Crystal structure of a decameric complex of human serum amyloid P component with bound dAMP. *Journal of Molecular Biology* **269**, 570–578.

HOSOKAWA, Y., MATSUMOTO, A., OKA, J., ITAKURA, H. AND YAMAGUCHI, K. (1990). Isolation and characterization of a cDNA for rat liver cysteine dioxygenase. *Biochemical and Biophysical Research Communications*. **168**, 473–478.

HOUGET, J., MEYER, A. AND PLANTEFOL, L. (1927). *Comptes Rendus de l'Academie des Sciences* **185**, 304.

HURKMAN, W.J., FORNARI, C.S. AND TANAKA, C.K. (1989). A comparison of the effect of salt on polypeptides and translatable mRNAs of a salt-tolerant and a salt-sensitive cultivar of barley. *Plant Physiology* **90**, 1444–1456.

HURKMAN, W.J., LANE, B.G. AND TANAKA, C.K. (1994). Nucleotide sequence of a transcript encoding a germin-like protein that is present in salt-stressed barley (*Hordeum vulgare*) roots. *Plant Physiology* **104**, 803–804.

HURKMAN, W.J., TAO, H.P. AND TANAKA, C.K. (1991). Germin-like polypeptides increase in barley roots during salt stress. *Plant Physiology* **97**, 366–374.

HURKMAN, W.J. AND TANAKA, C.K. (1996a). Effect of salt stress on germin gene expression in barley roots. *Plant Physiology* **110**, 971–977.

HURKMAN, W.J. AND TANAKA, C.K. (1996b). Germin gene expression is induced in wheat leaves by powdery mildew infection. *Plant Physiology* **111**, 735–739.

INAMDAR, K.V., TARACHAND, U., DEVASAGAYAM, T.P.A. AND RAGHAVAN, K.G. (1986). Immune complex of banana oxalate oxidase: use in quantitation of urinary oxalate. *Analytical Letters* **19**, 1987–1999.

INAMDAR, K.V., TARACHAND, U. AND RAGHAVAN, K.G. (1989). Determination of urinary oxalate with disposable oxalate oxidase membrane strips. *Analytical Letters* **22**, 841–851.

JAIKARAN, A.S.I., KENNEDY, T.D., DRATEWKA-KOS, E. AND LANE, B.G. (1990). Covalently bonded and adventitious glycans in germin. *Journal of Biological Chemistry* **265**, 12503–12512.

JENDROSSEK, D., FRISSE, A., BEHRENDS, A., ANDERMANN, M., KRATZIN, H.D., STANILAWSKI, T. AND SCHLEGEL, H.G. (1995). Biochemical and molecular characterisation of the *Pseudomonas lemoignei* polyhydroxyalkanoate depolymerase. *Journal of Bacteriology* **177**, 596–607.

JIANG, X.M., NEAL, B., SANTIAGO, F., LEE, S.J., ROMANA, L.K. AND REEVES, P.R. (1991). Structure and sequence of the *rfb* (O antigen) gene cluster of *Salmonella* serovar *typhimurium* (strain LT2). *Molecular Microbiology* **5**, 695–713.

JONES, A.M. (1994). Auxin-binding proteins. *Annual Review of Plant Physiology and Plant Molecular Biology*. **45**, 393–420.

KANEKO T., SATO, S., KOTANI, H., TANAKA, A., ASAMIZU, E., NAKAMURA, Y., MIYAJIMA, N., HIROSAWA, M., SIGIURA, M., SASAMOTO, S., KIMURA, T., HOSOUCHI, T., MATSUNO, A., MURAKI, A., NAKASAKI, N., NARUO, K., OKUMURA, S., SHIMPO, S., TAKEUCHI, C., WADA, T., WATANABE, A., YAMADA, M., YASUDA, M. AND TABATA, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Research* **3**, 109–136.

KOIKE, K., YOSHINO, K., SUE, N., UMEHARA, Y., ASHIKAWA, I., KURATA, N., AND SASAKI,T. (1997). Physical mapping of rice chromosomes 4 and 7 using YAC clones. *DNA Research* **4**, 27–33.

KOPLIN, R., WANG, G., HOTTE, B., PRIEFER, U.B. AND PUHLER, A. (1993). A 3.9-kb DNA region of *Xanthomonas campestris* pv. *campestris* that is necessary for lipopolysaccharide production encodes a set of enzymes involved in the synthesis of dTDP-rhamnose. *Journal of Bacteriology* **175**, 7782–7792.

KOTSIRA, V.P. AND CLONIS, Y.D. (1997). Oxalate oxidase from barley roots: purification to homogeneity and study of some molecular, catalytic, and binding properties. *Archives of Biochemistry and Biophysics* **340**, 239–249.

KOYAMA, H. (1988). Purification and characterization of oxalate oxidase from *Pseudomonas* sp. OX-53. *Agricultural and Biological Chemistry* **52**, 743–748.

KRUGEL, H., SCHUMANN, G., HANEL, F. AND FIEDLER, G. (1993). Nucleotide sequence analysis of five putative *Streptomyces griseus* genes, one of which complements an early function in daunorubicin biosynthesis that is linked to a putative gene cluster involved in TDP-daunosamine formation. *Molecular and General Genetics* **241**, 193–202.

KUCHHAL, N.K., SATYPAL AND PUNDIR, C.S. (1993). Partial purification, some properties and possible physiological role of an oxalate oxidase from grain sorghum leaves. *Indian Journal of Plant Physiology* **36**, 159–162.

KUHNER, C.H., HARTMAN, P.A. AND ALLISON, M.J. (1996). Generation of a proton motive force by the anaerobic oxalate-degrading bacterium *Oxalobacter formigenes*. *Applied and Environmental Microbiology* **62**, 2494–2500.

KURATA, N., NAGAMURA, Y., YAMAMOTO, K., HARUSHIMA, Y., SUE, N., WU, J, ANTONIO, B.A., SHOMURA, A., SHIMUZU, T., YIN, S-Y., INOUE, T., FUKUDA, A., SHIMANO, T., KUBOKI, Y., TOYAMA, T., MIYAMOTO, Y., KIRIHARA, T., HAYASAKA, K., MIYAO, A., MONNA, L., ZHONG, H.S., TAMURA, Y., WANG, Z-X., MOMMA, T., UMEHARA, Y., YANO, M., SASAKI, T. AND MINOBE, Y. (1994). A 300 kilobase interval genetic map of rice including 883 expressed sequence tags. *Nature Genetics* **8**, 365–372.

LABROU, N.E. AND CLONIS, Y.D. (1995). Biomimetic dye-ligands for oxalate recognizing enzymes. Studies with oxalate oxidase and oxalate decarboxylase. *Journal of Biotechnology* **40**, 59–70.

LAGUDAH, E.S., APELS, R. AND BROWN, A.D.H. (1991). The molecular-genetic analysis of *Triticum tauschii*, the D genome donor to hexaploid wheat. *Genome* **36**, 913–918.

LAKER, M.F., HOFMAN, A.F. AND MEEUSE, B.J. (1980). Spectrophotometric determination of urinary oxalate with oxalate oxidase prepared from moss. *Clinical Chemistry* **26**, 827–830.

LANE, B.G. (1991). Cellular desiccation and hydration: developmentally regulated proteins, and the maturation and germination of seed embryos. *FASEB Journal* **5**, 2893–2901.

LANE, B.G. (1994). Oxalate, germin, and the extracellular matrix of higher plants. *FASEB Journal* **8**, 294–301.

LANE, B.G., BERNIER, F., DRATEWKA-KOS, E., SHAFEI, R., KENNEDY, T.D., PYNE, C., MUNRO,

J.R., VAUGHAN, T., WALTERS, D. AND ALTOMARE, F. (1991). Homologies between members of the germin gene family in hexaploid wheat and similarities between these wheat germins and certain *Physarum* spherulins. *Journal of Biological Chemistry* **266**, 10461–10468.

LANE, B.G., CUMING, A.C., FRÉGEAU, J., CARPITA, N.C., HURKMAN, W.J., BERNIER, F., DRATEWKA-KOS, E. AND KENNEDY, T.D. (1992). Germin isoforms are discrete temporal markers of wheat development. *European Journal of Biochemistry* **209**, 961–969.

LANE, B.G., DUNWELL, J.M., RAY, J.A., SCHMITT, M.R. AND CUMING, A.C. (1993). Germin, a marker of early plant development, is an oxalate oxidase. *Journal of Biological Chemistry* **268**, 12239–12242.

LANE, B.G., GRZELCZAK, Z.F., KENNEDY, T.D., KAJIIOKA, R., ORR, J., D'AGOSTINO, S. AND JAIKARAN, A. (1986). Germin: Compartmentation of two forms of the protein by washing growing wheat embryos. *Biochemistry and Cell Biology* **64**, 1025–1037.

LATHIKA, K.M., INAMDAR, K.V., TARACHAND, U., SINGH, B.B. AND RAGHAVAN, K.G. (1995a). Determination of urinary oxalate using banana oxalate oxidase: comparison with immobilized enzyme. *Analytical Letters* **28**, 425–442.

LATHIKA, K.M., SHARMA, S., INAMBAR, K.V. AND RAGHAVAN, K.G. (1995b). Oxalate depletion from leafy vegetables using alginate entrapped banana oxalate oxidase. *Biotechnology Letters* **17**, 407–410.

LAWRENCE, M.C., IZARD, T., BEUCHAT, M., BLAGROVE, R.J. AND COLMAN, P.M. (1994). Structure of phaseolin at 2.2 Å resolution: Implications for a common vicilin/legumin structure and the genetic engineering of seed storage proteins. *Journal of Molecular Biology* **238**, 748–776.

LAZARUS, C.M. AND MACDONALD, H. (1996). Characterization of a strawberry gene for auxin-binding protein, and its expression in insect cells. *Plant Molecular Biology* **31**, 267–277.

LEBLANC, N., ROUX, C., PRADIER, J.M. AND PERROT-RECHENMANN, C. (1997). Characterization of two cDNAs encoding auxin-binding proteins in *Nicotiana tabacum*. *Plant Molecular Biology* **33**, 679–689.

LEHEL, C., LOS, D.A., WADA, H., GYORGYEI, J., HORVATH, I., KOVACS, E., MURATA, N. AND VIGH, L. (1993). A second *groEL*-like gene, organised in a *groESL* operon is present in the genome of *Synechocystis* sp. PCC 6803. *Journal of Biological Chemistry* **268**, 1799–1804.

LI, Y. AND STROHL, W.R. (1996). Cloning, purification,and properties of a phosphotyrosine protein from *Streptomyces coelicolor* A3(2). *Journal of Bacteriology* **178**, 136–142.

LIBERT, B. AND FRANCESCHI, V.R. (1987). Oxalate in crop plants. *Journal of Agricultural and Food Chemistry* **35**, 926–938.

LOEWUS, F.A., SAITO, K., SUTO, R.K. AND MARING, E. (1995). Conversion of D-arabinose to D-erythroascorbic acid and oxalic acid in *Sclerotinia sclerotiorum*. *Biochemical and Biophysical Research Communications* **212**, 196–203.

LUNG, H-Y., CORNELIUS, J.G. AND PECK, A.B. (1991). Cloning and expression of the oxalyl-CoA decarboxylase gene from the bacterium, *Oxalobacter formigenes*: Prospects for gene therapy to control Ca-oxalate kidney stone formation. *American Journal of Kidney Diseases* **17**, 381–385.

LUNG, H-Y., BAETZ, A.L. AND PECK, A.B. (1994). Molecular cloning, DNA sequence, and gene expression of the oxalyl-coenzyme A decarboxylase gene, *oxc*, from the bacterium *Oxalobacter formigenes*. *Journal of Bacteriology* **176**, 2468–2472.

LUPATELLI, F., PEDRAZZINI, E., BOLLINI, R., VITALE, A. AND CERIOTTI, A. (1997). The rate of phaseolin assembly is controlled by the glucosylation state of its N-linked oligosaccharide chains. *The Plant Cell* **9**, 597–609.

MACPHERSON, D.F., MANNING, P.A. AND MORONA, R. (1994). Characterisation of the dTDP-rhamnose biosynthetic genes encoded in the *rfb* locus of *Shigella flexneri*. *Molecular Microbiology* **11**, 281–292.

MALHERBE, P., KOHLER, C., DA PRADA, M., LANG, G., KIEFER, V., SCHWARCZ, R., LAHM, H.W. AND CESURA, A.M. (1994). Molecular cloning and functional expression of human 3-hydroxyanthranilic-acid dioxygenase. *Journal of Biological Chemistry* **269**, 13792–13797.

MARCIANO, P., DI LENNA, P. AND MAGRO, P. (1983). Oxalic acid, cell wall degrading enzymes

and their significance in the virulence of two *Sclerotinia sclerotiorum* isolates on sunflower. *Plant Pathology* **22**, 339–345.

MAROLDA, C.L. AND VALVANO, M.A. (1995). Genetic analysis of the dTDP-rhamnose biosynthesis region of *Escherichia coli* VW187 (O7:K1) rfb gene cluster: identification of functional homologs of *rfbB* and *rfbA* in the rff cluster and correct location of *rffe* gene. *Journal of Bacteriology* **177**, 5539–5546.

MASIREVIC, S. AND GULYA, T.J. (1992). Sclerotinia and Phomopsis-two devastating sunflower pathogens. *Field Crops Research* **30**, 271–300.

MAY, T.B., SHINABARGER, D., BOYD, A. AND CHAKRABARTY, A.M. (1994). Identification of amino acid residues involved in the activity of phosphomannose isomerase-guanosine 5'-diphospho-D-mannose pyrophosphorylase. A bifunctional enzyme in the alginate biosynthetic pathway of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* **269**, 4872–4877.

MCCUBBIN, W.C., CYRIL, M.K., KENNEDY, T.D. AND LANE, B.G. (1987). Germin: physico-chemical properties of the glycoprotein which signals the onset of growth in the germinating wheat embryo. *Biochemistry and Cell Biology* **65**, 1039–1048.

MEEUSE, B.J.D. AND CAMPBELL, J.M. (1959). An inhibitor of oxalic acid oxidase in beet extracts. *Plant Physiology* **34**, 583–586.

MEHTA, A. AND DATTA, A. (1991). Oxalate decarboxylase from *Collybia velutipes*. Purification, characterisation and cloning. *Journal of Biological Chemistry* **266**, 23548–23553.

MICALES, J.A. (1995). Oxalate decarboxylase in the brown-rot wood decay fungus *Postia placenta*. *Material Organismen* **29**, 177–186.

MICHALOWSKI, C.B. AND BOHNERT, H.J. (1992). Nucleotide sequence of a root-specific transcript encoding a germin-like protein from the halophyte *Mesembryanthemum crystallinum*. *Plant Physiology* **100**, 537–538.

MILLNER, P.A. (1995). The auxin signal. *Current Opinions in Cell Biology* **7**, 224–231.

MITCHISON, M., BULACH, D.M., VINH, T., RAJAKUMAR, K., FAINE, S. AND ADLER, B. (1997). Identification and characterisation of the dTDP-rhamnose biosynthesis and transfer genes of the lipopolysaccharide-related *rfb* locus in *Leptospira interrogans* serovar Copenhageni. *Journal of Bacteriology* **179**, 1262–1267.

MORTON, R.A., DILWORTH, M.J. AND WIENECKE, B. (1994). Biological degradation of sodium oxalate with bacillus agal N91/005579. US Patent 5314806.

MURZIN, A.G. (1993). Can homologous proteins evolve different enzymatic activities? *Trends in Biochemical Science* **18**, 403–405.

NAPIER, R.M. (1995). Towards an understanding of ABP1. *Journal of Experimental Botany* **46**, 1787–1795.

NAPIER, R.M. AND VENIS, M.A. (1995). Auxin action and auxin-binding proteins. *New Phytologist* **129**, 167–201.

NELSON, E.K. AND HASSELBRING, H. (1931). Some organic acids of wheat plants. *Journal of the American Chemical Society* **53**, 1040–1043.

NELSON, E.K. AND MOTTERN, H.H. (1931). Some organic acids in barley, maize, oats and rye plants. *Journal of the American Chemical Society* **53**, 3046–3048.

NOYES, R.D. AND HANCOCK, J.G. (1981). Role of oxalic acid in the *Sclerotinia* wilt of sunflower. *Physiological Plant Pathology* **18**, 123–132.

OBZANSKY, D.M. AND RICHARDSON, K.E. (1983). Quantification of urinary oxalate with oxalate oxidase from beet stems. *Clinical Chemistry* **29**, 1815–1819.

OBZANSKY, D.M. AND RICHARDSON, K.E. (1984). Oxalate oxidase composition for assay of oxalate. US Patent 4455371.

OH, H. AND CHOI, Y. (1994). Sequence analysis of β-xylosidase gene from *Bacillus stearothermophilus*. *Korean Journal of Applied Microbiology and Biotechnology* **22**, 134–142.

OHMIYA, A., KIKUCHI, M., SAKAI, S. AND HAYASHI, T. (1993). Purification and properties of an auxin-binding protein from the shoot apex of peach tree. *Plant Cell Physiology* **34**, 177–183.

OLSEN, C. (1939). Absorption of calcium and oxalic acid in higher green plants. *Comptes Rendus de Travaux Laboratoire Carlsberg* **23**, 101–124.

OLSEN, H.S., COLEMAN, T.A. AND ADAMS, M.D. (1997). Human oxalyl-CoA decarboxylase. US Patent 5635616.

ONO, M., SAGE-ONO, K., YASUI, M., OKASAKI, M. AND HARADA, H. (1993). Changes in polypeptides in Pharbitis cotyledons during the first flower-inductive period. *Plant Science* **89**, 135–145.

ONO, M., SAGE-ONO, K., INOUE, M., KAMADA, H. AND HARADA, H. (1996). Transient increase in the level of mRNA for a germin-like protein in leaves of the short-day plant *Pharbitis nil* during the photoperiodic induction of flowering. *Plant Cell Physiology* **37**, 855–861.

PALLADIN, V.I. AND LOVCHINOVSKAYA, E.I. (1916). Decomposition of oxalic acid by plants (in Russian). *Bulletin of the Academy of Science, Petrograd*, 937–948. Abstract in *Chemical Abstracts* (1917) **11**, 48.

PALME, K., HESSE, T., CAMPOS, N., GARBERS, C. AND YANOFSKY, M.F. (1992). Molecular analysis of an auxin binding protein gene located on chromosome 4 of Arabidopsis. *The Plant Cell* **4**, 193–201.

PEPKOWITZ, L.P., GILBERT, S.G. AND SHIVE, J.W. (1941). The importance of oxygen in the nutrient substrate for plants-organic acid. *Soil Science* **58**, 295–303.

PETERING, H., HAMMERSCHMIDT, S., FROSCH, M., VAN PUTTEN, J.P., ISON, C.A. AND ROBERTSON, B.D. (1996). Genes associated with meningococcal capsule complex are also found in *Neisseria gonorrhoeae*. *Journal of Bacteriology* **178**, 3342–3345.

PETRARULO, M., CERELLI, E., MARANGELLA, M., COSSEDDU, D., VITALE, C. AND LINARI, F. (1994). Assay of plasma oxalate with soluble oxalate oxidase. *Clinical Chemistry* **40**, 2030–2034.

PHILIPP, W.J., POULET, S., EIGLMEIER, K., PASCOPELLA, L., BALASUBRAMANIAN, V., HEYM, B., BERGH, S., BLOOM, B.R., JACOBS, W.R. AND COLE, S.T. (1996). An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proceedings of the National Academy of Sciences USA* **93**, 3132–3137.

PISSOWOTZKI, K., MANSOURI, K. AND PIEPERSBERG, W. (1991). Genetics of streptomycin production in *Streptomyces griseus*: molecular structure and putative function of genes strELMB2N. *Molecular and General Genetics* **231**, 113–123.

PRANKERD, T.L. (1920). Statocytes of the wheat haulm. *Botanical Gazette* **70**, 148–152.

PROUDFOOT, A.E.I., TURCATTI, G., WELLS, T.N.C., PAYTON, M.A. AND SMITH, D.J. (1994). Purification, cDNA cloning and heterologous expression of human phosphomannose isomerase. *European Journal of Biochemistry* **219**, 415–423.

PROUDFOOT, A.E., GOFFIN, L., PAYTON, M.A., WELLS, T.N. AND BERNARD, A.R. (1996). *In vivo* and *in vitro* folding of a recombinant metalloenzyme, phosphomannose isomerase. *Biochemical Journal* **318**, 437–442.

PUNDIR, C.S. (1991). Purification and properties of an oxalate oxidase from *Sorghum* leaves. *Phytochemistry* **30**, 1065–1067.

PUNDIR, C.S. AND KUCHHAL, N.K. (1989). Detection of an oxalate oxidase in grain sorghum roots. *Phytochemistry* **28**, 2909–2912.

PUNDIR, C.S. AND NATH, R. (1984). Occurrence of an oxalate oxidase in *Sorghum* leaves. *Phytochemistry* **23**, 1871–1874.

RAGHAVAN, K.G. AND TARACHAND, U. (1986). Degradation of oxalate in rats implanted with immobilized oxalate oxidase. *FEBS Letters* **195**, 101–105.

RENN, D. (1997). Biotechnology and the red seaweed polysaccharide industry: status, needs and prospects. *Trends in Biotechnology* **15**, 9–14.

ROBERTSON, B.D., FROSCH, M. AND VAN PUTTEN, J.P. (1994). The identification of cryptic rhamnose biosynthesis genes in *Neisseria gonorrhoeae* and their relationship to lipopolysaccharide biosynthesis. *Molecular Microbiology* **11**, 281–292.

ROWE, D.E. (1993). Oxalic acid effects in exudates of *Sclerotinia trifolium* and *S. sclerotiorum* and potential uses in selection. *Crop Science* **33**, 1146–1149.

RUML, L.A., PEARLE, M.S. AND PAK, C.Y.C. (1997). Medical therapy: calcium oxalate urolithiasis. *Urologic Clinics of North America* **24**, 117–133.

SACKSTON, W.E. (1992). On a treadmill: breeding sunflowers for resistance to disease. *Annual Review of Phytopathology* **30**, 529–551.

SAKA AMINI, M.A. AND VALLON, J.J. (1994). Comparison of performance and analytical applications of two immobilized oxalate sensors. *Analytica Chimica Acta* **299**, 75–79.

SAU, S. AND LEE, C.Y. (1996). Cloning of type 8 capsule genes and analysis of gene clusters of different capsular polysaccharides in *Staphylococcus aureus*. *Journal of Bacteriology* **178**, 2118–2126.

SCHWOB, E., CHOI, S-Y., SIMMONS, C., MIGLIACCIO, F., ILAG, L., HESSE, T., PALME, K. AND SÖLL, D. (1993). Molecular analysis of three maize 22kDa auxin-binding protein genes – transient expression and regulatory regions. *The Plant Journal* **4**, 423–432.

SELKOV, E., MALTSEV, N., OLSEN, G.J., OVERBEEK, R., WOESE, C.R. AND WHITMAN, W.B. (1997). A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. Available at:-http://www.cme.msu.edu/WIT/Doc/mj_recon.html

SHIMAZONO, H. (1955). Oxalic acid decarboxylase, a new enzyme from the mycelia of wood destroying fungi. *Journal of Biochemistry* **42**, 321–340.

SHIMOKAWA, K., UEDA, Y. AND KASAI, Z. (1972). Decarboxylation of oxalic acid during ripening of banana fruit (*Musa sapientum L.*). *Agricultural and Biological Chem*istry **36**, 2021–2024.

SHIMOMURA, S., LIU, W., INOHARA, N., WATANABE, S. AND FUTAI, M. (1993). Structure of the gene for an auxin-binding protein and a gene for 7SL RNA from *Arabidopsis thaliana*. *Plant Cell Physiology* **34**, 633–637.

SIDHU, H., OGDEN, S.D., LUNG, H.Y., LUTTGE, B.G., BAETZ, A.L. AND PECK, A.B. (1997). DNA sequencing and expression of the formyl coenzyme A transferase gene, *frc*, from *Oxalobacter formigenes*. *Journal of Microbiology* **179**, 3378–3381.

SLIGHTOM, J.L., DRONG, R.F., KLASSEY, R.C. AND HOFFMAN, L.M. (1985). Nucleotide sequences from phaseolin cDNA clones: The major storage proteins from *Phaseolus vulgaris* are encoded by two unique gene families. *Nucleic Acids Research* **13**, 6483–6498.

SMITH, C.A., PINKNEY, M., GUINEY, D.G. AND THOMAS, C.M. (1993). The ancestral IncP replication system consisted of contiguous *oriV* and *trfA* segments as deduced from a comparison of the nucleotide sequences of diverse IncP plasmids. *Journal of General Microbiology* **139**, 1761–1766.

SMITH, J.J., THOMSON, A.J., PROUDFOOT, A.E. AND WELLS, T.N. (1997). Identification of an Fe(III)-dihydroxyphenylalanine site in recombinant phosphomannose isomerase from *Candida albicans*. *European Journal of Biochemistry* **244**, 325–333.

SPERRY, J.S. AND TYREE, M.T. (1988). Mechanism of water stress-induced xylem embolism. *Plant Physiology* **88**, 581–587.

SRINAVASAN, N., WHITE, H.E., EMSLEY, J., WOOD, S.P., PEPYS, M.B. AND BLUNDELL, T.L. (1994). Comparative analyses of pentraxins: implications for protomer assembly and ligand binding. *Structure* **2**, 1017–1027.

STAEHELIN, M. (1919). Die Rolle der Oxalsäure in der Pflanze. Enzymatische Abbau des Oxalations. *Biochemische Zeitung* **96**, 1–49.

STEVENSON, G., NEAL, B., LIU, HOBBS, M., PACKER, N.H., BATLEY, M., REDMOND, J.W., LINDQUIST, L. AND REEVES, P. (1994). Structure of the O antigen of *Escherichia coli* L-12 and the sequence of its *rfb* gene cluster. *Journal of Bacteriology* **176**, 4144–4156.

STOKKERMANS, J.P.W.G., PIERIK, A.J., WOLBERT, R.B.G., HAGEN, W.R., VAN DONGEN, W.M.A.M. AND VEEGER, C. (1992). The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from *Desulfovibrio vulgaris* (Hildenborough). *European Journal of Biochemistry* **208**, 435–442.

SUGIURA, M., YAMAMURA, H., HIRANO,K., SASAKI, M., MORIKAWA, M. AND TSUBOI, M. (1979). Purification and properties of an oxalate oxidase from barley seedlings. *Chemical and Pharmaceutical Bulletin* **27**, 2003–2007.

SWART, S., LOGMAN, T.J.J., SMIT, G., LUGTENBERG, B.J.J. AND KIJNE, J.W. (1994). Purification and partial characterisation of a glycoprotein from pea (*Pisum sativum*) with receptor activity for rhicadhesin, an attachment protein of Rhizobiaceae. *Plant Molecular Biology* **24**, 171–183.

THOMPSON, C., DUNWELL, J.M., JOHNSTONE, C.E., LAY, V., RAY, J., WATSON, H. AND NISBET, G. (1995a). Introduction of genes for oxalate degrading enzymes into plants as a potential means of providing tolerance to *Sclerotinia*. *IOBC wprs Bulletin* **18**, 34–38.

THOMPSON, C., DUNWELL, J.M., JOHNSTONE, C.E., LAY, V., RAY, J., SCHMITT, M., WATSON, H. AND NISBET, G. (1995b). Degradation of oxalic acid by transgenic oilseed rape plants expressing oxalate oxidase. *Euphytica* **85**, 169–172.

THORDAHL-CHRISTENSEN, H., ZHANG, Z., WEI, Y. AND COLLINGE, D.B. (1997). Subcellular localization of H_2O_2 in plants. H_2O_2 accumulation in papillae and hypersensitive response during the powdery-mildew interaction. *The Plant Journal* **11**, 1187–1194.

THORSON, J.S., LO, S.F., PLOUX, O., HE, X. AND LIU, H.W. (1994). Studies on the biosynthesis of 3,6-dideoxyhexoses: molecular cloning and characterization of the *asc* (ascarylose) region from *Yersinia pseudotuberculosis* serogroup VA. *Journal of Bacteriology* **176**, 5483–5493.

TRINCHANT, J-C., GUÉRIN, V. AND RIGAUD, J. (1994). Acetylene reduction by symbiosomes and free bacteroids from broad bean (*Vicia faba* L.) nodules: role of oxalate. *Plant Physiology* **105**, 555–561.

TRINCHANT, J-C. AND RIGAUD, J. (1996). Bacteroid oxalate oxidase and soluble oxalate in nodules of faba beans (*Vicia faba* L.) submitted to water restricted conditions: possible involvement in nitrogen fixation. *Journal of Experimental Botany* **47**, 1865–1870.

VARALAKSHMI, P., LATHIKA, K.M., RAGHAVAN, K.G. AND SINGH, B.B. (1995). Altered physicochemical characteristics of polyethylene glycol linked beet stem oxalate oxidase. *Biotechnology and Bioengineering* **46**, 254–257.

VARALAKSHMI, P. AND RICHARDSON, K.E. (1992). Studies on oxalate oxidase from beet stems upon immobilisation on concanavalin A. *Biochemical International* **26**, 153–162.

VENIS, M.A. AND NAPIER, R. (1995). Auxin receptors and auxin binding proteins. *Critical Reviews in Plant Science* **14**, 27–47.

VITALE, A., BIELLI, A. AND CERIOTTI, A. (1995). The binding protein associates with monomeric phaseolin. *Plant Physiology* **107**, 1411–1418.

WADLEIGH, C.H. AND SHIVE, J.W. (1939). Organic acid content of corn plants as influenced by pH of substrate and form of nitrogen supplied. *American Journal of Botany* **26**, 244–248.

WAGNER, G.J. (1981). Vacuolar deposition of ascorbate-derived oxalic acid in barley. *Plant Physiology* **67**, 591–593.

WANG, L., ROMANNA, L.K. AND REEVES, P.R. (1992). Molecular analysis of a *Salmonella enterica* group E1 *rfb* gene cluster: O antigen and the genetic basis for the major polymorphism. *Genetics* **130**, 429–443.

WEI, Y.D. (1994). The biochemical and molecular basis of interactions between barley and the powdery mildew fungus. PhD Thesis. The Royal Veterinary and Agricultural University, Frederiksberg, Denmark.

WOJTASZEK, P. (1997). Oxidative burst: an early plant response to pathogen infection. *Biochemical Journal* **322**, 681–692.

WORLEY, K.C., WIESE, B.A. AND SMITH, R.F. (1990). BEAUTY: an enhanced BLAST-based search tool that integrates multiple biological information resources into sequence similarity search results. *Genome Research* **5**, 173–184.

WYMAN, H. AND PALMER, J.K. (1964). Organic acids in the ripening banana fruit. *Plant Physiology* **39**, 630–633.

YAMATO, S., WAKABAYASHI, H., NAKAJIMA, M. AND SHIMADA, K. (1994). Amperometric determination of oxalate in plasma and urine by liquid chromatography with immobilized oxalate oxidase. *Journal of Chromatography* **656**, 29–35.

YAMAZAKI, M., THORNE, L., MIKOLAJCZAK, M., ARMENTROUT, R.W. AND POLLOCK, T.J. (1996). Linkage of genes essential for synthesis of a polysaccharide capsule in *Sphingomonas* S88. *Journal of Bacteriology* **178**, 2676–2687.

YAO, Z. AND VALVANO, M.A. (1994). Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotypes Y and 4a. *Journal of Bacteriology* **176**, 4133–4143.

YOHDA, M., ENDO, I., ABE, Y., OHTA, T., IIDA, T., MARUYAMA, T. AND KAGAWA, Y. (1996). Gene for aspartate racemase from the sulfur-dependent hyperthermophilic achaeum, *Desulfurococcus* strain SY. *Journal of Biological Chemistry* **271**, 22017–22021.

YOSHIDA, K., SEKI, S., FUJIMURA, M., MIWA, Y. AND FUJITA, Y. (1995). Cloning and

sequencing of a 36-kb region of the *Bacillus subtilis* genome between the *gnt* and *iol* operons. *DNA Research* **2**, 61–69.

YOSHIHARA, T., SOGAWA, K., PATHAK, M.D., JULIANO, B.O. AND SAKAMURA, S. (1990). Oxalic acid as a sucking inhibitor of the brown planthopper in rice (Delphacidae, Homoptera). *Entomology, Experimental and Applied* **27**, 149–155.

ZALESKI, W. AND KUKHARKOVA, A.M. (1928). Enzymic oxidation of oxalic acid by higher plants.(in Russian). *Ukraine Khemical Zhurnal* **3**, 139–152. Abstract in *Chemical Abstracts* (1929) **23**, 2991.

ZALESKI, W. AND REINHARD, A. (1912). Über die fermentative Oxydation der Oxalsäure. *Biochemische Zeitung* **33**, 449–455.

ZHANG, L., AL-HENDY, A., TOIVANEN, P. AND SKURNIK, M. (1993). Genetic organisation and sequence of the *rfb* gene cluster of *Yersinia enterocolitica* serotype O:3: similarities to the dTDP-L-rhamnose biosynthesis pathway of *Salmonella* and to the bacterial polysaccharide transport systems. *Molecular Microbiology* **9**, 309–321.

ZHANG, Z., COLLINGE, D.B. AND THORDAHL-CHRISTENSEN, H. (1995). Germin-like oxalate oxidase, a H_2O_2-producing enzyme, accumulates in barley attacked by the powdery mildew fungus. *The Plant Journal* **8**, 139–145.

ZHANG, Z., YANG, J., COLLINGE, D.B. AND THORDAHL-CHRISTENSEN, H. (1996). Ethanol increases sensitivity of oxalate oxidase assays and facilitates direct activity staining in SDS gels. *Plant Molecular Biology Reporter* **14**, 266–272.